AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXXII
1950

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXII, Number 1
JANUARY, 1950

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

The AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications should be sent to Professor AUREL WINTNER at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at specia rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

m
c
o
i

w
(
t
b

I
(
a

SIMPLE HOMOTOPY TYPES.*

By J. H. C. WHITEHEAD.

1. Introduction. This is a sequel to two papers [1] entitled "Combinatorial Homotopy," Parts (I) and (II). It deals with what I have previously called the "nucleus," but which will now be called the *simple homotopy type* of a complex. It is closely related to parts of [1] and [3] but the treatment is so different that we shall start again from the beginning.

Let $\{K\}$ be the class of all (cell) complexes,[2] as defined in CH (I), which are of the same homotopy type as a given complex K. Let $K' \equiv K$ (i. e. $K' \varepsilon \{K\}$) and let $\bar{\phi}: K \equiv K'$ be the class of maps which are homotopic to a given homotopy equivalence, $\phi: K \equiv K'$. If $\phi': K' \equiv K''$, we define $\bar{\phi}'\bar{\phi}$ by

$$\bar{\phi}'\bar{\phi} = \overline{\phi'\phi}: K \equiv K''.$$

It is easily verified that the classes ϕ, with this multiplication, form a groupoid,[3] G, whose unit elements are the classes $\bar{1}: K' \equiv K'$, for every $K' \varepsilon \{K\}$, where $1: K' \to K'$ is the identical map. Our plan is to analyse this groupoid in algebraic terms.

First consider the group, $G_K \subset G$, which consists of the classes $\bar{\phi}: K \equiv K$. We define an additive Abelian group, T, which depends only on $\pi_1(K)$. The group T admits G_K as a group of operators and we shall define a crossed homomorphism $\tau: G_K \to \mathrm{T}$. We call $\tau(g)$ the torsion of a given element $g \varepsilon G_K$. If $\phi: K \equiv K'$, where $K' \neq K$, we define a class of elements $\tau(\phi) \subset \mathrm{T}$, which we call the *torsion* of ϕ. We describe ϕ as a *simple* (*homotopy*) *equivalence* if, and only if, $\tau(\phi) = 0$. We say that K and K' are of the same *simple homotopy type,* and shall write $K \equiv K'\ (\Sigma)$, if, and only if, there is a simple equivalence $\phi: K \equiv K'$. It will follow from the

* Received January 18, 1949.

[1] *Bulletin of the American Mathematical Society,* vol. 55 (1949), pp. 213-45 and 453-96. These papers will be referred to as CH (I) and CH (II).

[2] Until the final section we assume that any given complex is finite and connected. We also assume that the points in our complexes are taken from some aggregate, σ, which is given in advance. The power of σ shall exceed that of the continuum, so that it is not exhausted by any one (finite) complex, and the points in Hilbert space shall be included in σ.

[3] See [6], p. 132.

definition of $\tau(\phi)$ that $K \equiv K' (\Sigma)$ is an equivalence relation. We then prove that $K \equiv K' (\Sigma)$ if, and only if, K can be transformed into K' by a "formal deformation," which is defined in much the same way as in [1]. Thus the elementary transformations, or "moves," do not appear in the definition of simple equivalence but in a theorem which is analogous to Tietze's theorem [4] on discrete groups. Similarly it is proved that two complexes are of the same n-type if, and only if, they can be interchanged by elementary transformations of the sort used in [1] to define the "n-group."

It was proved in [3] that the Reidemeister-Franz torsion,[5] when defined, is an invariant of the simple homotopy type. Using this fact, examples were given of complexes, which are of the same homotopy type but not of the same simple homotopy type. However, if $\mathrm{T} = 0$, then $K \equiv K' (\Sigma)$ if $K \equiv K'$. It will be obvious that this is so if $\pi_1(K) = 1$. It follows from Theorems 14, 15 in [11] that $\mathrm{T} = 0$ if $\pi_1(K)$ is of order 2, 3, 4 or cyclic infinite.

It is an open question whether or not the simple homotopy type is a topological invariant. However we shall prove that it is a combinatorial invariant in the following sense. If K' is a sub-division of K, then the identical map $K \to K'$ is a simple equivalence.[6] Any differentiable manifold has a "preferred" class of triangulations,[7] any two of which are combinatorially equivalent in the sense of Newman. Also any analytic variety has a preferred class of triangulations,[7] any two of which have a common sub-division. Therefore the simple homotopy type has an invariant status in differential and alegbraic geometry and in the study of analytic varieties.

2. The group T. Let R be a ring with a unit element 1. Eventually R will be the group ring [8] of $\pi_1(K)$ but here we only assume that, if A is a free R-module of (finite) rank n, then any free R-module, which is isomorphic [9] to A, also has rank n. This condition is equivalent to the

---

[4] See [7], p. 46.

[5] See [8], [9] and p. 1209 of [3]. In Section **12** below it is shown, in the case of Lens space, how this is related to our torsion. See also [10].

[6] This may turn out to be a wider definition, even for simplicial complexes, than the one based on Newman's "moves," or on recti-linear sub-divisions (see [12], [13]). For example, we do not enquire whether or not the vertex scheme of a given "curvilinear" triangulation of an n-simplex is a formal n-element, as defined by Newman.

[7] See [2] and [14].

[8] By the group ring of a group, Γ, we shall always mean the integral group ring, in which the additive group is the ordinary free Abelian group, which is freely generated by the elements of Γ.

[9] A module will always mean a free R-module and, unless the contrary is stated, a homomorphism will always mean an operator homomorphism.

condition that every regular R-matrix (i. e. one with elements in R and a 2-sided inverse) is square. Hence it is satisfied if there is a homomorphism, other than $R \to 0$, of R into a division ring, D. For such a homomorphism carries a regular R-matrix into a regular D-matrix, which is necessarily square. If R is the group ring of a group, Γ, then $\Gamma \to 1$ defines a homomorphism of R into the rational field. Therefore the condition of rank invariance is satisfied.

Let M be the module, of infinite rank, whose elements are the infinite sequences $(r_1, r_2, \cdots)$ $(r_i \varepsilon R)$, in which all but a finite number of $r_1, r_2, \cdots$ are zero. The elements in R will operate on M from the left.[10] Thus an operator, $r \varepsilon R$, transforms m into rm, where

$$m = (r_1, r_2, \cdots), \qquad rm = (rr_1, rr_2, \cdots).$$

Let $m_i \varepsilon M$ be the basis element which is given by $r_i = 1$, $r_j = 0$ if $j \neq i$. Let $M^n \subset M$ be the module generated[11] by $(m_1, \cdots, m_n)$ and M_n the one generated by $(m_{n+1}, m_{n+2}, \cdots)$, where $n \geq 0$ and $M^0 = 0$. Then M is the direct sum $M = M^n + M_n$ and a given element in M is in M^n for some value of n. We shall describe an endomorphism, $f: M \to M$, as *admissible* if, and only if, $fm_i = m_i$ for all sufficiently large values of i. If $f, g: M \to M$ are admissible endomorphisms[12] so, obviously, is $fg: M \to M$ and if $f: M \to M$ is an admissible automorphism so is f^{-1}. Therefore the admissible automorphims form a group, $\mathfrak{a}$.

Let $f: M \to M$ be an (admissible) endomorphism and let $fm_j = m_j$ if $j > p$. Let n_i be such that $fm_i \varepsilon M^{n_i}$ $(i = 1, \cdots, p)$ and let $n \geq \operatorname{Max}(n_i, p)$. Then $fm_i \varepsilon M^n$ for $i = 1, \cdots, n$ and $fm_j = m_j$ if $j > n$. Therefore $fM^n \subset M^n$, $fm = m$ if $m \varepsilon M_n$. We shall write $f = (f)^n: M \to M$, and f^n will denote the endomorphism, $f^n: M^n \to M^n$, which is induced by f. That is to say, $f^n m = fm$ if $m \varepsilon M^n$. Notice that $(f)^n = (f)^q$ if $q > n$. Therefore, if $f_i: M \to M$ is any finite set of endomorphisms, we may take $f_i = (f_i)^n$, for any value of n which is sufficiently large to be the same for each i. Notice also that any endomorphism $f': M^n \to M^n$ can be extended to a unique endomorphism, $(f)^n: M \to M$, such that $f^n = f'$. Obviously f^n is an automorphism if, and only if, $f \varepsilon \mathfrak{a}$.

Let $f = (f)^n$ be given by

$$fm_i = \sum_{j=1}^{\infty} f_{ij} m_j \qquad (f_{ij} \varepsilon R). \tag{2.1}$$

---

[10] This has the disadvantage indicated by (2.3) below. But the convention $m \to mr$ would be inconvenient in the geometrical application.

[11] I. e. generated with the help of the operators in R.

[12] Unless the contrary is stated it is to be assumed that any given endomorphism of M is admissible.

Then the matrix $\boldsymbol{f} = [f_{ij}]$ is of the form

$$\boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}^n & 0 \\ 0 & \mathbf{1}_\infty \end{bmatrix} \tag{2.2}$$

where $\boldsymbol{f}^n$ is the matrix of $f^n : M^n \to M^n$ and $\mathbf{1}_\infty$ is the infinite unit matrix. Let $g: M \to M$ be given by

$$gm_i = \sum_j g_{ij} m_j.$$

Since $fr = rf$, where $r \varepsilon R$ is any operator, we have

$$fgm_i = \sum_j g_{ij} f m_j = \sum_{j,k} g_{ij} f_{jk} m_k. \tag{2.3}$$

Therefore $fg: M \to M$ corresponds to the matrix $\boldsymbol{gf}$.

Let $g: M \to M$ be given by

$$gm_i = m_i + rm_j, \qquad gm_k = m_k \qquad (j, k \neq i; r \varepsilon R). \tag{2.4}$$

Then g has an inverse, which is given by (2.4), with r replaced by $-r$. It is therefore an (admissible) automorphism. Let $\Sigma_1 \subset \mathfrak{A}$ be the group generated by all such automorphisms, for all values of i, j, r.

Let A and B be the modules generated by disjoint sub-sets, $m_{i_1}, \cdots, m_{i_p}$ and $m_{j_1}, \cdots, m_{j_q}$, of the basis elements $m_1, m_2, \cdots,$. Let $h: A \to B$ be an arbitrary homomorphism and let $g: M \to M$ be given by

$$g(a + b) = a + (ha + b), \qquad gm_l = m_l, \tag{2.5}$$

where $a \varepsilon A$, $b \varepsilon B$ and $l \neq i_\rho$ or j_σ. Then g is the resultant of the homomorphisms

$$m_{i_\rho} \to m_{i_\rho} + h_{\rho\sigma} m_{j_\sigma}, \qquad m_k \to m_k \qquad (k \neq i_\rho),$$

where $hm_{i_\rho} = h_{\rho 1} m_{j_1} + \cdots + h_{\rho q} m_{j_q}$. These are of the form (2.4), whence $g \varepsilon \Sigma_1$.

THEOREM 1. *Σ_1 is an invariant sub-group of $\mathfrak{A}$ and $\mathfrak{A}/\Sigma_1$ is Abelian.*

Let $f, f' \subset \mathfrak{A}$. We shall write $f \equiv f'$ if, and only if, $f = gf'g'$, where $g, g' \subset \Sigma_1$. This is obviously an equivalence relation. Assume that $ff' \equiv f'f$ for every pair $f, f' \subset \mathfrak{A}$. Let $g \varepsilon \Sigma_1$, and let $f' = gf^{-1}$. Then

$$fgf^{-1} \equiv gf^{-1}f = g.$$

Therefore $fgf^{-1} \varepsilon \Sigma_1$, whence Σ_1 is invariant in $\mathfrak{A}$. Also $\mathfrak{A}/\Sigma_1$ is Abelian since $ff' \equiv f'f$ for every pair $f, f' \subset \mathfrak{A}$.

We proceed to prove that $ff' \equiv f'f$. Let $A = M^p$, let B be the module

generated by $m_{p+1}, \cdots, m_{2p}$ and let g be given by (2.5). Then $g = (g)^{2p}$ and

$$g^{2p} = \begin{bmatrix} \mathbf{1}_p & \boldsymbol{h} \\ 0 & \mathbf{1}_p \end{bmatrix},$$

where $\boldsymbol{h} = [h_{\rho\sigma}]$ and $\mathbf{1}_p$ is the unit matrix of order p. Let $f = (f)^{2p} \varepsilon\, \mathfrak{A}$ and let

$$\boldsymbol{f}^{2p} = \begin{bmatrix} \boldsymbol{f}_{11} & \boldsymbol{f}_{12} \\ \boldsymbol{f}_{21} & \boldsymbol{f}_{22} \end{bmatrix},$$

where $\boldsymbol{f}_{11}, \boldsymbol{f}_{22}$ are square matrices of order p. Let $\boldsymbol{f}'^{2p}$ and $\boldsymbol{f}'_{\lambda\mu}$ $(\lambda, \mu = 1, 2)$ be similarly defined in terms of $f' = (f')^{2p}$. We shall write $\boldsymbol{f}^{2p} \equiv \boldsymbol{f}'^{2p}$ if, and only if, $f \equiv f'$. Then

$$\boldsymbol{f}^{2p} \equiv \boldsymbol{f}^{2p}\boldsymbol{g}^{2p} = \begin{bmatrix} \boldsymbol{f}_{11} & \boldsymbol{f}_{11}\boldsymbol{h} + \boldsymbol{f}_{12} \\ \boldsymbol{f}_{21} & \boldsymbol{f}_{21}\boldsymbol{h} + \boldsymbol{f}_{22} \end{bmatrix}. \tag{2.6}$$

Similarly a right hand multiple of the second column may be added to the first. Also $\boldsymbol{f}^{2p} \equiv \boldsymbol{g}^{2p}\boldsymbol{f}^{2p}$, and $\boldsymbol{g}^{2p}\boldsymbol{f}^{2p}$ is obtained from $\boldsymbol{f}^{2p}$ by a similar operation on the rows.

Let $f, f' \subset \mathfrak{A}$ be given and let p be so large that

$$f = (f)^p = (f)^{2p}, \qquad f' = (f')^p = (f')^{2p}.$$

Let $\boldsymbol{r} = \boldsymbol{f}^p$, $\boldsymbol{r}' = \boldsymbol{f}'^p$. Then $\boldsymbol{r}, \boldsymbol{r}'$ are regular matrices. Therefore, beginning with (2.6), with $\boldsymbol{h} = \boldsymbol{r}^{-1}$ and $\boldsymbol{f}$ replaced by $\boldsymbol{f}'\boldsymbol{f}$, we have

$$\boldsymbol{f}'^{2p}\boldsymbol{f}^{2p} = \begin{bmatrix} \boldsymbol{r}'\boldsymbol{r} & 0 \\ 0 & \mathbf{1}_p \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{r}'\boldsymbol{r} & \boldsymbol{r}' \\ 0 & \mathbf{1}_p \end{bmatrix} \equiv \begin{bmatrix} 0 & \boldsymbol{r}' \\ -\boldsymbol{r} & \mathbf{1}_p \end{bmatrix}$$
$$\equiv \begin{bmatrix} 0 & \boldsymbol{r}' \\ -\boldsymbol{r} & 0 \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{r} & \boldsymbol{r}' \\ -\boldsymbol{r} & 0 \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{r} & \boldsymbol{r}' \\ 0 & \boldsymbol{r}' \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{r} & 0 \\ 0 & \boldsymbol{r}' \end{bmatrix}.$$

Similarly

$$\begin{bmatrix} 0 & \boldsymbol{r}' \\ -\boldsymbol{r} & 0 \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{r}' & 0 \\ 0 & \boldsymbol{r} \end{bmatrix} \equiv \boldsymbol{f}^{2p}\boldsymbol{f}'^{2p}.$$

Therefore $\boldsymbol{f}\boldsymbol{f}' \equiv \boldsymbol{f}'\boldsymbol{f}$ and the theorem is proved.

Since $\mathfrak{A}/\Sigma_1$ is Abelian it follows that $\mathfrak{A}^c \subset \Sigma_1$, where $\mathfrak{A}^c$ is the commutator sub-group of $\mathfrak{A}$. Therefore we have the corollary:

COROLLARY. *If $\Sigma \subset \mathfrak{A}$ is any sub-group, which contains Σ_1, then Σ is invariant and $\mathfrak{A}/\Sigma$ is Abelian.*

The totality of automorphisms $(f)^n \varepsilon\, \mathfrak{A}$, for a fixed value of n, is obviously a sub-group, $(\mathfrak{A})^n \subset \mathfrak{A}$. It follows from Theorem 1 that

$(\Sigma_1)^n = \Sigma_1 \cap (\mathfrak{A})^n$ is an invariant sub-group [13] of $(\mathfrak{A})^n$ and that $(\mathfrak{A})^n/(\Sigma_1)^n$ is Abelian. Let $\mathfrak{A}^n$ be the group of (operator) automorphisms, $f^n : M^n \to M^n$, and let $\phi : (\mathfrak{A})^n \to \mathfrak{A}^n$ be given by $\phi(f)^n = f^n$. Then ϕ is obviously an isomorphism.[14] It follows from the invariance of $(\Sigma_1)^n$ in $(\mathfrak{A})^n$ that $\Sigma_1{}^n = \phi(\Sigma_1)^n$ is invariant in $\mathfrak{A}^n$ and that $\Sigma_1{}^n$ is independent of the particular isomorphism $\phi : (\mathfrak{A})^n \approx \mathfrak{A}^n$. Also $\mathfrak{A}^n/\Sigma_1{}^n$ is Abelian.

Let Λ be a sub-group of the multiplicative group of regular elements in R (that is, elements with two-sided inverses), which contains both ± 1. Let $g : M \to M$ be given by

$$(2.7) \qquad gm_i = \lambda m_i + rm_j, \qquad gm_k = m_k \qquad (j, k \neq i),$$

where $\lambda \varepsilon \Lambda$, $r \varepsilon R$. Then $g \varepsilon \mathfrak{A}$ and g^{-1} is given by (2.7) with λ, r replaced by λ^{-1}, $-\lambda^{-1}r$. Let Σ_Λ be the sub-group of $\mathfrak{A}$, which is generated by all automorphisms of the form (2.7), for every choice of i, j, λ and r. Clearly $\Sigma_1 \subset \Sigma_\Lambda$. Therefore Σ_Λ is invariant and $\mathrm{T} = \mathfrak{A}/\Sigma_\Lambda$ is Abelian. We shall keep Λ fixed and shall write Σ, T for Σ_Λ, T_Λ. The elements of Σ will be called *simple automorphisms*. We shall write T additively and $\tau(f) \varepsilon \mathrm{T}$ will denote the co-set containing a given $f \varepsilon \mathfrak{A}$.

Our "torsion" will be defined in terms of T. An element of torsion will correspond to an isomorphism of one module, of finite rank, onto another. In order to classify such isomorphisms in term of T we need a standard class, which have "zero torsion." We therefore proceed to define a class of "basic modules" in M, which are related by a standard class of automorphisms, called permutations.

By a *basic module*, $A \subset M$, we shall mean the one generated by $m_{i_1}, \cdots, m_{i_p}$, for any (distinct) values of $i_1, \cdots, i_p$. We shall call $(m_{i_1}, \cdots, m_{i_p})$ *the* basis of A. We allow $p = 0$, in which case the set $(m_{i_1}, \cdots, m_{i_p})$ is empty and $A = M^0$. Let $p \geq 0$ and let M_A be the module generated by the remaining basic elements, $m_j \neq m_{i_\rho}$, of M. Then M is the direct sum $M = A + M_A$. Let B be a basic module and let $(m_{j_1}, \cdots, m_{j_q})$ be its basis. We shall only allow ourselves to form the direct sum $A + B = B + A$, if $A \cap B = 0$. In this case $A + B$ will be the basic module, whose basis is

$$(m_{i_1}, \cdots, m_{i_p}, m_{j_1}, \cdots, m_{j_q}),$$

---

[13] The example II, on p. 1233 of [3] shows that $(\Sigma_1)^n$ may be a larger group than the one which is generated by transformations of the form (2.4), with $i, j \leq n$. I see no reason to suppose that the latter is necessarily an invariant sub-group of $(a)^n$.

[14] An isomorphism, without qualification, will always mean an isomorphism onto.

not the set of all pairs (a, b), with $a \varepsilon A$, $b \varepsilon B$. Let C be a given basic module. Then $C = A + B$ will always mean that A, B are basic modules, with disjoint bases, of which C is the direct sum.

Let $i_1, \cdots, i_n$ be any permutation of $1, \cdots, n$, for any $n \geqq 1$. Let $P: M \to M$ be the automorphism, which is given by

$$Pm_j = m_{i_j}, \qquad Pm_k = m_k \qquad (j = 1, \cdots, n; k > n).$$

We shall call P a *permutation*. It follows from (2.7), with $\lambda, r = \pm 1$, that the transformations

$$(m_i, m_j) \to (-m_i + m_j, m_j) \to (-m_i + m_j, m_i) \to (m_j, m_i)$$

determine simple automorphisms. Therefore $P \varepsilon \Sigma$. Let A, B be basic modules of the same rank and let n be so large that the bases of A, B are both contained in M^n. Then there is obviously a permutation, $P = (P)^n$, such that $PA = B$. The totality of permutations is obviously a sub-group of $\mathfrak{a}$.

Let $\alpha: A \approx A'$, where A, A' are basic modules. Since A and A' have the same rank, accoording to our condition on R, there is a permutation, P, such that $PA' = A$. Let $f: M \to M$ be given by

$$(2.8) \qquad f(a + m) = P\alpha a + m \qquad (a \varepsilon A, m \varepsilon M_A),$$

and let $\tau(\alpha) = \tau(f)$. Let P' be any other permutation such that $P'A' = A$ and let f' be defined by (2.8), with P replaced by P'. Since $P'P^{-1}A = A$ the permutation $P'P^{-1}$ permutes the basis elements of A among themselves. Therefore $P'': M \to M$, given by

$$P''(a + m) = P'P^{-1}a + m \qquad (m \varepsilon M_A),$$

is a permutation. Since $P\alpha A = A$, it follows from (2.8) that $f' = P''f$. Therefore

$$\tau(f') = \tau(P'') + \tau(f) = \tau(f).$$

Therefore $\tau(\alpha)$ does not depend on the choice of P. We shall describe α as a *simple isomorphism*, and shall write $\alpha: A \approx A'$ (Σ), if, and only if, $\tau(\alpha) = 0$. It follows from (2.8) that $\tau(\alpha) = 0$ if $\alpha = 1: A \approx A$. In particular $\tau(\alpha) = 0$ if $A = A' = M^0$.

Let α, P, f, mean the same as in (2.8), let $\alpha': A' \approx A''$ and let P' be a permutation such that $P'A'' = A'$. Then $\tau(\alpha') = \tau(f')$, $\tau(\alpha'\alpha) = \tau(f'')$, where f' and f'' are given by (2.8) with α, P replaced by α', P' and by $\alpha'\alpha$, PP'. Clearly $P^{-1}M_A = M_{A'}$. Therefore

$$Pf'P^{-1}f(a+m) = Pf'P^{-1}(P\alpha a + m) = Pf'(\alpha a + P^{-1}m)$$
$$= P(P'\alpha'\alpha a + P^{-1}m) = PP'\alpha'\alpha a + m = f''(a+m).$$

Therefore

$$\tau(\alpha'\alpha) = \tau(Pf'P^{-1}f) = \tau(f') + \tau(f) = \tau(\alpha') + \tau(\alpha).$$

Since $\tau(1) = 0$ it follows that $\tau(\alpha^{-1}) = -\tau(\alpha)$.

Let $\alpha : A \approx A'$ be the isomorphism induced by a permutation, $P' : M \to M$, such that $P'A = A'$. Then f, given by (2. 8), is a permutation. Therefore $\tau(\alpha) = 0$.

Let A, B and A', B' be two pairs of basic modules such that $A \cap B = A' \cap B' = 0$. Let $\gamma : A + B \to A' + B'$ be a homomorphism such that $\gamma B \subset B'$. Then $\gamma(a + b) = \alpha a + (ha + \beta b)$, where $\alpha a \,\varepsilon\, A', ha, \beta b \subset B'$. It is easily verified that $\alpha : A \to A'$, $h : A \to B'$, $\beta : B \to B'$ are homomorphisms.

THEOREM 2. *If either:*

(i) γ *is an isomorphism*[14] *and either* α *is an isomorphism into or* β *is onto, or if*

(ii) α, β *are isomorphisms, then* α, β, γ *are all isomorphisms and*

$$\tau(\gamma) = \tau(\alpha) + \tau(\beta).$$

Let $\gamma : A + B \approx A' + B'$. If $\beta b = 0$, then $\gamma b = \beta b = 0$, whence $b = 0$. Therefore β is an isomorphism into. Let $a' \,\varepsilon\, A'$ be given. Then $\alpha a + (ha + \beta b) = \gamma(a + b) = a'$ for some $a \,\varepsilon\, A$, $b \,\varepsilon\, B$. Since $ha + \beta b \,\varepsilon\, B'$ we have $\alpha a = a'$. Therefore α is onto. Let $\alpha : A \approx A'$ and let $b' \,\varepsilon\, B'$ be given. Then $\alpha a + (ha + \beta b) = b'$ for some $a \,\varepsilon\, A$, $b \,\varepsilon\, B$. Since $\alpha a \,\varepsilon\, A'$ we have $\alpha a = 0$. Therefore $a = 0$, $ha = 0$ and $\beta b = b'$. Therefore β is onto. Let $\beta : B \approx B'$ and let $\alpha a = 0$. Then $\gamma(a - \beta^{-1}ha) = \alpha a + (ha - ha) = 0$. Therefore $a - \beta^{-1}ha = 0$. Since $\beta^{-1}ha \,\varepsilon\, B$ it follows that $a = 0$. Therefore $\alpha : A \approx A'$. Thus α, β, γ are isomorphisms if (i) is satisfied.

Let α, β be isomorphisms. Then $\gamma = \gamma^*\delta$, where $\gamma^* : A + B \to A' + B'$, $\delta : A + B \to A + B$ are given by $\gamma^*(a + b) = \alpha a + \beta b$, $\delta(a + b) = a + (\beta^{-1}ha + b)$. Obviously γ^*, δ are isomorphisms and so therefore is γ. Moreover $g : M \to M$ is of the form (2. 5), where

$$g(a + b + m) = \delta(a + b) + m \qquad (m \,\varepsilon\, M_{A+B}).$$

Let P be a permutation such that $PA' = A$, $PB' = B$. Let $f = f_\alpha$ be defined by (2. 8) and let f_β, f_γ, f_{γ^*} be similarly defined in terms of β, γ, γ^* and the same permutation P. Then $f_\alpha b = b$, $f_\beta a = a$ and

$$f_\gamma(a + b) = P\alpha a + P(ha + \beta b) = f_\alpha + f_\beta(\beta^{-1}ha + b) = f_\alpha f_\beta g(a + b).$$

Since $g \varepsilon \Sigma$ it follows that

$$\tau(\gamma) = \tau(f_\gamma) = \tau(f_\alpha) + \tau(f_\beta) + \tau(g) = \tau(\alpha) + \tau(\beta)$$

and the proof is complete.

COROLLARY. *If any two of α, β, γ are simple isomorphisms, so is the third.*

Let $\theta: R \approx R$ be an automorphism of R and let $s_\theta: M \to M$ be the transformation which is given by

$$s_\theta(r_1, r_2, \cdots) = (\theta r_1, \theta r_2, \cdots). \tag{2.9}$$

Obviously $s_{\theta^{-1}} = s_\theta^{-1}$ and $s_\theta s_\phi = s_{\theta\phi}$, where $\phi: R \approx R$. Also $s_\theta(rm) = (\theta r) s_\theta m$, where $r \varepsilon R$, $m \varepsilon M$. Hence it follows that, if $f: M \to M$ is an (operator) endomorphism, then $(s_\theta f s_\phi) rm = (\theta\phi r)(s_\theta f s_\phi) m$. Therefore $f^\theta r = r f^\theta$ where $f^\theta = s_\theta f s_\theta^{-1}$. Since $s_\theta m_i = m_i$ it follows that $f^\theta \varepsilon \mathfrak{A}$ if $f \varepsilon \mathfrak{A}$. Let $g: M \to M$ be given by (2.7). Since $s_\theta m = m_i$ we have $g^\theta m_i = (\theta\lambda) m_i + m_j$, $g^\theta m_k = m_k$. Therefore g^θ is also of the form (2.7) if $\theta\lambda \varepsilon \Lambda$. Clearly $(g_1 g_2)^\theta = g_1^\theta g_2^\theta$ and it follows that $f^\theta \varepsilon \Sigma$ if $f \varepsilon \Sigma$, provided $\theta\Lambda \subset \Lambda$.

We shall describe $\theta: R \approx R$ as a Λ-*automorphism* if, and only if, $\theta\Lambda = \Lambda$. The totality of Λ-automorphisms is obviously a group Θ. Since $f^\theta \varepsilon \Sigma$ if $f \varepsilon \Sigma$ and $\theta \varepsilon \Theta$ it follows that T admits Θ as a group of operators, according to the rule

$$\theta\tau(f) = \tau(f^\theta). \tag{2.10}$$

Let $x \varepsilon R$ be any regular element, not necessarily an element of Λ, and let $\theta_x r = x r x^{-1}$. I say that

$$\theta_x \tau = \tau \tag{2.11}$$

for each $\tau \varepsilon$ T. For let $f \varepsilon \mathfrak{A}$ be given by (2.1). Then

$$f^{\theta_x} m_i = s_{\theta_x} \Sigma_j f_{ij} m_j = \Sigma_j (x f_{ij} x^{-1}) m_j.$$

Let $f = (f)^n$ and let $g_x = (g_x)^n: M \to M$ be given by

$$g_x(r_1, \cdots, r_n, r_{n+1}, \cdots) = (r_1 x, \cdots, r_n x, r_{n+1}, \cdots).$$

Then $g_x m_i = x m_i$ if $i \leq n$. Since $f x m = x f m$,

$$f g_x m_i = x f m_i = \sum_{j=1}^n x f_{ij} m_j = \sum_{j=1}^n (x f_{ij} x^{-1}) x m_j = g_x f^{\theta_x} m_i \quad (i = 1, \cdots, n).$$

Therefore $f^{\theta_x} = g_x^{-1} f g_x$ and $\tau(f^{\theta_x}) = -\tau(g_x) + \tau(f) + \tau(g_x) = \tau(f)$, which proves (2.11).

Let $f \varepsilon \mathfrak{A}$ and let $\boldsymbol{f}$ and $\boldsymbol{f}^n$ mean the same as in (2.2). Let g be given by (2.7) and let $\boldsymbol{g}$ be its matrix. Then $\boldsymbol{gf}$ is obtained from $\boldsymbol{f}$ by the following operations

(2.12)
a) *multiplying a row from the left by an element* $\lambda \varepsilon \Lambda$,
b) *adding a left multiple of one row to another, the multiplier being an arbitrary element* $r \varepsilon R$.

Therefore $f \varepsilon \Sigma$ if, and only if, $\boldsymbol{f} \rightarrow \boldsymbol{1}_\infty$ by a finite sequence of such transformations. Let $\boldsymbol{f} \rightarrow \boldsymbol{1}_\infty$ by such a sequence, $\sigma_1, \cdots, \sigma_p$, and let $\boldsymbol{f} = (f)^n$. Then there is a $k \geq 0$ such that no row of $\boldsymbol{f}$, after the $(n+k)$-th is involved in any of $\sigma_1, \cdots, \sigma_p$. Therefore $\sigma_1, \cdots, \sigma_p$ transform [13] $\boldsymbol{f}^{n+k}$ into $\boldsymbol{1}_{n+k}$, where

$$\boldsymbol{f}^{n+k} = \begin{bmatrix} \boldsymbol{f}^n & 0 \\ 0 & \boldsymbol{1}_k \end{bmatrix}.$$

Let R be the group ring of a group Γ and let Λ consist of the elements $\pm\gamma$, where $\gamma \varepsilon \Gamma$. If Γ is Abelian, the determinant, $|\boldsymbol{f}^n|$, of $\boldsymbol{f}^n$ can be calculated in the ordinary way. Obviously $|\boldsymbol{f}^n|$ is unaltered by (2.12b) or by an "expansion," $\boldsymbol{f}^n \rightarrow \boldsymbol{f}^{n+k}$, and a transformation of the form (2.12a) changes $|\boldsymbol{f}^n|$ into $\pm\gamma\,|\boldsymbol{f}^n|$. Therefore $\pm|\boldsymbol{f}^n| \,\varepsilon\, \Gamma$ if $f \varepsilon \Sigma$. Let Γ be cyclic of order 5 and let $\gamma \neq 1$. Then $(1-\gamma-\gamma^4)(1-\gamma^2-\gamma^3)=1$. Therefore $f: M \rightarrow M$, given by $f(r_1, r_2, r_3, \cdots) = \{r_1(1-\gamma-\gamma^4), r_2, r_3, \cdots\}$ is in $\mathfrak{A}$, but not in Σ. Therefore $\mathrm{T} \neq 0$. On the other hand it follows from the theory of integral, unimodular matrices, in case $\Gamma = 1$, and from Theorems 14, 15 in [11], that $\mathrm{T} = 0$ if Γ is of order 1, 2, 3, 4 or is cyclic infinite.

We continue, until Section **9**, without the assumption that R is a group ring.

3. Chain systems. By a *chain system*, $C = \{C_n\}$, we shall mean a family of basic modules, $C_n \subset M$, together with a boundary operator, $\partial = \{\partial_n\}$, which is a family of (operator) homomorphisms, $\partial_n: C_n \rightarrow C_{n-1}$, such that $\partial_n \partial_{n+1} = 0$. For the sake of completeness we define $\partial_0 C_0 = C_{-1} = 0$. Each C_n, being a basic module, is of finite rank. We do not require C_0 to be of rank 1, as we did in section **8** of CH(II). For example, we allow $C_0 = 0$. We assume that $C_n = 0$ for all sufficiently large values of n. If $C_n = 0$ when $n > N \geq 0$, but $C_N \neq 0$, we write $N = \dim C$. We write $C = 0$, and dim $C = -1$, if $C_n = 0$ for every $n \geq 0$. We insist that $C_p \cap C_q = 0$ if $p \neq q$ and C shall be the set-theoretic union of the groups $C_0, C_1, \cdots$. Thus $c \,\varepsilon\, C$ means that $c \,\varepsilon\, C_n$ and $c + c'$ is only defined if $c, c' \subset C_n$, for some $n \geq 0$. Also ∂ is a map, $\partial: C \rightarrow C$, of the set C into itself.

Until Section 9 we shall only consider chain mappings,[15] $f: C \to C'$, of C into a chain system, $C' = \{C'_n\}$, such that each $f: C_n \to C'_n$ is an operator homomorphism. That is to say, in the terminology of CH (II), f is associated with the identical isomorphism, $R \to R$. Thus[16] $\partial f = f\partial$, $fr = rf$, where $r \varepsilon R$ is an operator. Also $f \simeq g: C \to C'$ will mean that

$$g_n - f_n = \partial_{n+1}\eta_{n+1} + \eta_n\partial_n \qquad (n \geqq 0), \tag{3.1}$$

where $\eta = \{\eta_n\}$ is a chain deformation operator, and $f: C \equiv. C'$ will mean that there is a chain mapping, $f': C' \to C$, such that $f'f \simeq 1$, $ff' \simeq 1$ in the sense of (3.1). We shall call $f: C \to C'$ a simple isomorphism, and shall write $f: C \approx C'$ (Σ), if, and only if, it is a chain mapping such that $f_n: C_n \approx C'_n$ (Σ), for each $n \geq 0$.

Let B, C be given chain systems and let $B_n = B'_n + B''_n$, $C_n = C'_n + C''_n$, where, according to our convention, B'_n, B''_n, C'_n, C''_n are basic modules. Let $f: B \to C$ be a chain mapping such that

$$f_n(b' + b'') = f'_n b' + (g_n b' + f''_n b'') \qquad (b' \varepsilon B'_n, b'' \varepsilon B''_n),$$

for each $n \geq 0$, where $f'_n b' \varepsilon C'_n$, $g_n b', f''_n b'' \subset C''_n$.

LEMMA 1. *If any two of $\{f_n\}$, $\{f'_n\}$, $\{f''_n\}$ are families of simple isomorphisms, so is the third.*

This follows immediately from the corollary to Theorem 2.

Let $C_n = C'_n + C''_n$ and let $\partial C'_n \subset C'_{n-1}$ for each $n \geqq 0$. Let $C' = \{C'_n\}$ and let $\partial': C' \to C'$ be defined by $\partial' c' = \partial c'$. Then $\partial'\partial' c' = \partial\partial c' = 0$. Under these, and only these conditions, we shall describe C', with the boundary operator ∂', as a *sub-system* of C. If also $\partial C''_n \subset C''_{n-1}$ $(n \geq 0)$, so that $\partial(c' + c'') = \partial' c' + \partial'' c''$ $(c' \varepsilon C'_n, c'' \varepsilon C''_n)$, then $C'' = \{C''_n\}$, with boundary operator ∂'', is also a sub-system. In this case we shall call C the *direct sum*, $C = C' + C'' = C'' + C'$, of C' and C''. Let C', C'' be given, disjoint,[17] chain systems. Then the direct sum, $C' + C''$, will be the system which consists of the groups $C'_n + C''_n$, with $\partial(c' + c'') = \partial' c' + \partial'' c''$. Similarly we define the direct sum of any finite set of disjoint chain systems.

---

[15] At this stage we do not impose any restriction such as $f_0 m_i = \lambda m_j$ on f_0, where m_i is a basis element of C_0. For example, $C \to 0$ is a chain mapping.

[16] We shall often use ∂ to denote the boundary operator in each of two or more systems, $C, C', C'', \cdots$, which occur in the same context. On other occasions we shall use $\partial, \partial', \partial'', \cdots$ to denote the boundary operators in $C, C, C'', \cdots$.

[17] We describe two or more basic modules, or chain systems, as disjoint if, and only if, $0 \varepsilon M$ is their only common element.

Let C' be a sub-system of C and let $C_n = C'_n + C''_n$. Let $j_n : C_n \to C''_n$ and $\partial''_n : C''_n \to C''_{n-1}$ be defined by

$$(3.2) \qquad j_n(c' + c'') = c'', \qquad \partial''_n c'' = j_{n-1}\partial_n c''.$$

Then $\partial''_n j_n = j_{n-1}\partial_n$, since $\partial C' \subset C'$ and $j_{n-1}C'_{n-1} = 0$. Therefore

$$\partial''_n\partial''_{n+1}j_{n+1} = \partial_n j_n \partial_{n+1} = j_{n-1}\partial_n\partial_{n+1} = 0,$$

whence $\partial''_n\partial''_{n+1} = 0$. Therefore $C'' = \{C''_n\}$, with $\partial'' = \{\partial''_n\}$ as boundary operator, is a chain system. We call it the *residue system*, mod C′, and write $C'' = C - C'$. Notice, however, that an element in C'' is an element in the basic module C''_n, for some $n \geq 0$, not a residue class of elements in C. Notice also that $j = \{j_n\}$ is a chain mapping, $j : C \to C''$; also that $c - jc \,\varepsilon\, C'$, whence

$$(3.3) \qquad \partial c'' - \partial'' c'' = \partial c'' - j\partial c'' \,\varepsilon\, C'.$$

Let B', C' be sub-systems of chain systems B, C and let $B'' = B - B'$, $C'' = C - C'$. Let $f : B \to C$ be a chain mapping such that $fB' \subset C'$. Then $f' : B' \to C'$, given by $f'b' = fb'$, is obviously a chain mapping. Let

$$f''_n = j_n f_n : B''_n \to C''_n.$$

Then $f''j = jf$, where $j : B \to B''$ is defined in the same way as $j : C \to C''$. Since $\partial''j = j\partial$ we have

$$f''\partial''j = jf\partial = j\partial f = \partial''f''j,$$

where ∂ operates on B, C and ∂'' on B'', C''. Therefore f'' is a chain mapping. We shall call $f' : B' \to C'$, $f'' : B'' \to C''$ the chain mappings *induced* by f. It follows from Lemma 1 that, if any two of f, f', f'' are simple isomorphisms, so is the third.

Let A be a common sub-system of B and C. Then we shall describe a chain mapping, $f : B \to C$, as *rel.* A if, and only if, $fa = a$ for each $a \,\varepsilon\, A$. We shall say that $f \simeq g : B \to C$, rel. A, if, and only if, $g - f = \partial\eta + \eta\partial$, where $\eta : B \to C$ is a deformation operator such that $\eta A = 0$.

Let $Z_n(C) = \partial_n^{-1}(0)$ and let

$$H_n(C) = Z_n(C) - \partial_{n+1}C_{n+1} \qquad (n \geq 0).$$

A chain mapping, $f : B \to C$, obviously induces a family of homomorphisms

$$f_* : H_n(B) \to H_n(C).$$

Let C' be a sub-system of C, let $i : C' \to C$ be the identical map, which is

obviously a chain mapping, and let $j: C \to C''$ mean the same as before. Let $z'' \varepsilon Z_n(C'')$. Then it follows from (3.3) that $\partial z'' \varepsilon Z_{n-1}(C')$. Therefore ∂ induces a family of homomorphisms $d_*: H_n(C'') \to H_{n-1}(C')$, where $H_{-1}(C') = 0$. It is known [18] that the sequence of homomorphisms,

$$(3.4) \qquad \cdots \xrightarrow[d_*]{} H_n(C') \xrightarrow[i_*]{} H_n(C) \xrightarrow[j_*]{} H_n(C'') \xrightarrow[d_*]{} \cdots \xrightarrow[d_*]{} H_{-1}(C'),$$

is exact, meaning that the kernel of each homomorphism is the image group of its predecessor. We prove that $d_* H_n(C'') = i_*^{-1}(0)$. Let $\bar{z} \varepsilon H_n(X)$ ($X = C, C'$ or C'') be the residue class containing a given element $z \varepsilon Z_n(X)$. Let $z'' \varepsilon Z_n(C'')$. Then $\partial z'' = i\partial z'' \varepsilon Z_{n-1}(C)$, and

$$i_* d_* \bar{z}'' = i_* \overline{\partial z''} = \overline{i\partial z''} = \overline{\partial z''} = 0.$$

Therefore $d_* H_n(C'') \subset i_*^{-1}(0)$. Conversely, let $i_* \bar{z} = 0$, where $z' \varepsilon Z_{n-1}(C')$. This means that $iz' \varepsilon \partial C_n$, or that $z' = \partial c = \partial(c' + z'')$ $(c' \varepsilon C'_n, z'' \varepsilon C''_n)$. Therefore, writing $z' - \partial c' = z'_1$, we have

$$\bar{z}' = \bar{z}'_1 = \overline{\partial z''} = d_* \bar{z}'',$$

whence $d_* H_n(C') = i_*^{-1}(0)$. It follows from similar arguments that $i_* H_n(C') = j_*^{-1}(0)$ and that $j_* H_n(C) = d_*^{-1}(0)$.

4. Deformation retracts. Let $C' \subset C$ be a sub-system and let $i: C \to C$ be the identical chain mapping. A chain mapping, $k: C \to C'$, will be called a *retraction* if, and only if, $ki = 1$. We shall call C' a *deformation retract* (D. R.) of C if, and only if, there is a retraction, $k: C \to C'$, such that $ik \simeq 1$, rel. C'. Let $ik \simeq 1$, rel. C', and let $k': C \to C'$ be any other retraction. Then $ik' \simeq ik'ik = ik \simeq 1$, rel. C'.

THEOREM 3. *A sub-system $C' \subset C$ is a D. R. of C if, and only if, $H_n(C - C') = 0$ for every $n \geq 0$.*

Let C' be a D. R. of C and let $k: C \to C'$ be a retraction. Then

$$(4.1) \qquad 1 - ik = \partial\eta + \eta\partial,$$

where $\eta: C \to C$ is a deformation operator such that $\eta C' = 0$. Let $C'' = C - C'$ and let $z'' \varepsilon Z_n(C'')$. Then $\partial z'' \varepsilon C'$, whence $\eta\partial z'' = 0$. Clearly $ji = 0$, $jz'' = z''$, where $j: C \to C''$ is given by (3.2). Therefore

$$z'' = j(1 - ik)z'' = j(\partial\eta + \eta\partial)z'' = j\partial\eta z'' = \partial j\eta z''.$$

Therefore $H_n(C'') = 0$ $(n \geq 0)$.

[18] See Theorem 3.3 in [15].

Conversely, let $H_n(C'') = 0$ for every $n \geq 0$. Assume that there are homomorphisms,

$$k_r : C_r \to C'_r,\ \eta_{r+1} : C_r \to C_{r+1} \qquad (r = -1, \cdots, n-1),$$

such that $k_r i_r = 1$, $\partial_r k_r = k_{r-1}\partial_r$ and

$$(4.2)_r \qquad i_r k_r - 1 = \partial_{r+1}\eta_{r+1} + \eta_r \partial_r,$$

these conditions being vacuous if $n = 0$. Let $m''_1, \cdots, m''_p$ ($m''_\lambda = m_\iota$) be the basis of C''_n and let

$$(4.3) \qquad (1 + \eta_n \partial_n) m''_\lambda = c'_\lambda + c''_\lambda \qquad (c'_\lambda \varepsilon C'_n, c''_\lambda = C''_n).$$

It follows from $(4.2)_{n-1}$ that

$$\partial_n(1 + \eta_n\partial_n) = (1 + \partial_n\eta_n)\partial_n = (i_{n-1}k_{n-1} - \eta_{n-1}\partial_{n-1})\partial_n = i_{n-1}k_{n-1}\partial_n.$$

Therefore $\partial_n(c'_\lambda + c''_\lambda) = i_{n-1}k_{n-1}\partial_n m''_\lambda = k_{n-1}\partial_n m''_\lambda$. Therefore $\partial''_n c''_\lambda = 0$. Since $H_n(C'') = 0$ we have $c''_\lambda = \partial''_{n+1} a''_\lambda$, for some $a''_\lambda \varepsilon C''_{n+1}$. Let $a'_\lambda = c''_\lambda - \partial_{n+1} a''_\lambda \varepsilon C'_n$. Then it follows from (4.3) that

$$(4.4) \qquad (1 + \eta_n\partial_n) m''_\lambda = c'_\lambda + a'_\lambda + \partial_{n+1} a''_\lambda.$$

Let $k_n : C_n \to C'_n$ and $\eta_{n+1} : C_n \to C_{n+1}$ be the operator homomorphisms defined by $k_n c' = c'$, $\eta_{n+1} c' = 0$ and

$$k_n m''_\lambda = c'_\lambda + a'_\lambda, \qquad \eta_{n+1} m''_\lambda = -a''_\lambda.$$

Then $(4.2)_n$ follows from (4.4). Also

$$\partial_n k_n c' = \partial_n c' = k_{n-1}\partial_n c'$$

$$\partial_n k_n m''_\lambda = \partial_n(c'_\lambda + a'_\lambda) = \partial_n(c'_\lambda + c''_\lambda) = k_{n-1}\partial_n m''_\lambda.$$

Therefore, starting with $k_{-1} = \eta_{-1} = 0$, the theorem follows by induction on n.

Corollary 1. *$C \equiv 0$ if, and only if, $H_n(C) = 0$ for every $n \geq 0$.*

Corollary. 2. *C' is a D.R. of C if, and only if, $C - C' \equiv 0$.*

Lemma 2. *If C' is a D.R. of C then $C \approx C' + C''$ (Σ), rel. C', where $C'' = C - C'$.*

Let $C^* = C' + C''$. Then $C^*_n = C_n$ and $\partial^* : C^* \to C^*$ is given by $\partial^*(c' + c'') = \partial' c' + \partial'' c''$. Let i, k, η mean the same as in (4.1) and let $f : C^* \to C$ be given by

$$f(c' + c'') = (c' - kc'') + c'' = c' + (\partial\eta + \eta\partial)c''.$$

Then $\partial fc' = \partial c' = f\partial^* c'$ and

$$\partial fc'' = \partial(\partial\eta + \eta\partial)c'' = \partial\eta\partial c'',$$
$$f\partial^* c'' = (\partial\eta + \eta\partial)\partial'' c'' = (\partial\eta + \eta\partial)(\partial c'' + c'_1) \qquad (c'_1 \varepsilon C').$$

Since $\partial C' \subset C'$ and $\eta C' = 0$ it follows that $f\partial^* c'' = \partial\eta\partial c'' = \partial fc''$. Therefore f is a chain mapping. It follows from Lemma 1 that $f: C^* \approx C\ (\Sigma)$ and the lemma is proved.

5. Simple equivalence. We shall describe a chain system, B, as *elementary* if, and only if, $B_n = 0$ when $n \neq r - 1$, r, for some $r \geqq 1$, and $\partial_r: B_r \approx B_{r-1}\ (\Sigma)$. This being so, it is obvious that $H_n(B) = 0$ for every $n \geqq 0$. Therefore $B \equiv 0$, by Theorem 3, Corollary 1. We shall describe B as *collapsible* if, and only if, it is the direct sum of a finite set of elementary systems. Clearly $B \equiv 0$ if B is collapsible. It is obvious that B' is collapsible if $B' \approx B\ (\Sigma)$, where B is collapsible; also that, if B, B' are disjoint and collapsible, then $B + B'$ is collapsible; also that the direct sum of a set of r-dimensional elementary systems is itself elementary.

Let $B_n = A_n + Z_n$, let $\partial'_n: A_n \approx Z_{n-1}\ (\Sigma)$ and let $\partial_n: B_n \to B_{n-1}$ be given by $\partial_n a = \partial'_n a$, $\partial_n Z_n = 0$, $(n = 1, 2, \cdots)$. Then $B = \{B_n\}$, with $\partial = \{\partial_n\}$ as boundary operator, is the direct sum of the elementary systems $(\cdots, 0, A_n, Z_{n-1}, 0, \cdots)$. Therefore B is collapsible and any collapsible system is obviously of this form.

We shall say that C, C' are in the same *simple equivalence class*, and shall write $C \equiv C'\ (\Sigma)$ if, and only if, there are collapsible systems, B, B', such that

$$f: B + C \approx B' + C'\ (\Sigma). \tag{5.1}$$

This being so, it follows from Theorem 3, Corollary 2, that C, C' are D. R.'s of $B + C$, $B' + C'$. Let

$$i: C \to B + C, \qquad i': C' \to B' + C'$$
$$k: B + C \to C, \qquad k': B' + C' \to C'$$

be the identity maps and any retractions. Let

$$g \simeq k'fi: C \to C' \tag{5.2}$$

and let $g' \simeq kf^{-1}i': C' \to C$. Then

$$g'g \simeq kf^{-1}i'k'fi \simeq kf^{-1}fi = 1. \tag{5.3}$$

Similarly $gg' \simeq 1$. Therefore $g: C \equiv C'$. We shall describe a chain mapping, $g: C \to C'$, as a *simple equivalence* and shall write $g: C \equiv C'$ (Σ), if, and only if, it is related by (5. 2) to some simple isomorphism of the type (5. 1). It follows from (5. 3) that, if $g: C \equiv C'$ (Σ) and if $g'': C' \to C$ is such that $gg'' \simeq 1$, then g'' is a simple equivalence. Obviously $g: C \equiv C'$ (Σ) if $g: C \approx C'$ (Σ).

The relation $C \equiv C'$ (Σ) is obviously reflexive and symmetric. We proceed to prove that it is transitive; also that, if $g: C \equiv C'$ (Σ) and $g': C' \equiv C''$ (Σ), then $g'g: C \equiv C''$ (Σ). Let C, C' be related by (5. 1), let $f^*: B^* + C' \approx B'' + C''$ (Σ), where B^*, B'' are collapsible, and assume that

$$(5.4) \qquad B' \cap B^* = B^* \cap (B + C) = B' \cap (B'' + C'') = 0.$$

Then $B + B^* + C \approx B' + B^* + C' \approx B' + B'' + C''$ (Σ), whence $C \equiv C''$ (Σ). If (5. 4) are not satisfied we apply a permutation, $P_n: B'_n \to A'_n$, to each module B'_n, thus transforming B' into a new system, A', such that $P: B' \approx A'$ (Σ), where $P = \{P_n\}$, and $A' \cap C = 0$. Let

$$h: B' + C' \approx A' + C' \ (\Sigma)$$

be given by $h(b' + c') = Pb' + c'$. Then

$$(5.5) \qquad hf: B + C \approx A' + C' \ (\Sigma).$$

Similarly we can replace B^* by $A^* = P^*B^*$. We can choose the basic modules A'_n, A^*_n in such a way that (5. 4) are satisfied when B', B^* are replaced by A', A^*. Therefore $C \equiv C''$ (Σ), and it follows that $C \equiv C'$ (Σ) is an equivalence relation.

Let $g: C \to C'$ be given by (5. 2) and let h mean the same as in (5. 5). Then $g \simeq (k'h^{-1})(hf)i: C \to C'$ and $k'h^{-1}: A' + C' \to C'$ is obviously a retraction. Also $k'h^{-1}$ can be extended to a retraction, $A' + A^* + C' \to C'$, by mapping A^* on zero. Therefore, if $g: C \equiv C'$ (Σ) and $g': C' \equiv C''$ (Σ) we lose no generality in assuming that g satisfies (5. 2) and that

$$f': B' + C' \approx B'' + C'' \ (\Sigma), \qquad g \simeq k''f'i': C' \to C'',$$

where $k'': B'' + C'' \to C''$ is a retraction. This being so,

$$f'f: B + C \approx B'' + C'' \ (\Sigma)$$

and

$$g'g \simeq k''f'i'k'fi \simeq k''f'fi: C \to C''.$$

Therefore $g'g: C \equiv C''$ (Σ).

A non-zero element, which is common to two chain systems, will be called an *accidental intersection*, unless it is in a common sub-system, which is explicitly mentioned in the context. Accidental intersections between any finite set of systems, $C, \cdots$, can always be eliminated, as in the paragraph containing (5. 5), by replacing $C, \cdots$, by a set of chain systems, $PC, \cdots$, where $P_n: C_n \to (PC)_n$, is a suitable set of permutations. When the context requires it, we shall always assume that this has already been done.

Let $C \equiv C'$ (Σ), $C^* \equiv C'^*$ (Σ). Then $B + C \approx B' + C'$ (Σ), $B^* + C^* \approx B'^* + C'^*$ (Σ), where B, B' etc. are collapsible. Therefore, in the absence of accidental intersections, it follows from Lemma 1, in Section **3**, that

$$B + B^* + C + C^* \approx B' + B'^* + C' + C'^* \ (\Sigma),$$

whence

$$C + C^* \equiv C' + C'^* \ (\Sigma). \tag{5.6}$$

Let A be a common sub-system of C and C'. We shall write $C \equiv C''$ (Σ), rel. A, if, and only if,

$$f: B + C \approx B' + C' \ (\Sigma), \text{rel. } A, \tag{5.7}$$

where B, B' are collapsible.

THEOREM 4. a) *If* $C \equiv C'$ (Σ), *rel.* A, *then* $C - A \equiv C' - A'$ (Σ).

b) *If* $C - A \equiv 0$ (Σ), *then* $C \equiv A$ (Σ), *rel.* A.

Let C, C' be related by (5. 7). Since f induces the identity, $A \to A$ (Σ), it follows from Lemma 1 that

$$f': (B + C) - A \approx (B' + C') - A \ (\Sigma),$$

where f' is the chain mapping induced by f. Obviously

$$(B + C) - A = B + (C - A), \qquad (B' + C') - A = B' + (C' - A),$$

which proves (a).

Let $C'' \equiv 0$ (Σ), where $C'' = C - A$. Then $B + C'' \equiv B'$ (Σ), where B, B' are collapsible. Since $C'' \equiv 0$ it follows from Theorem 3, Corollary 2, that A is a D. R. of C. Therefore $C \approx A + C''$ (Σ), rel. A, by Lemma 2. Therefore

$$B + C \approx B + C'' + A \approx B' + A \ (\Sigma), \text{rel. } A.$$

Therefore $C \equiv A$ (Σ), rel. A, and the theorem is proved.

By a (p, q)-system, C, we shall mean a chain system such that $C_n = 0$ if $n < p$ or if $n > q$ $(p \leq q)$.

2

LEMMA [19] 3. *If $C \equiv C'$ (Σ), where C, C' are (p,q)-systems, there are collapsible (p,q)-systems, B, B', such that $B + C \approx B' + C'$ (Σ).*

Let $f : A + C \approx A' + C'$ (Σ), where A, A' are collapsible. Let $n = \dim A > q$. Then $n = \dim(A + C) = \dim(A' + C')$. It follows from the definition of a collapsible system that

$$A = B^1 + \cdots + B^m, \tag{5.8}$$

where each B^i is an elementary system. Let E be the direct sum of all the n-dimensional summands, B^i, and let D be the direct sum of the others. Let $D', E' \subset A'$ be similarly defined. Then $E_n = (A + C)_n$, $E'_n = (A' + C')_n$, and $\partial_n : E_n \approx E_{n-1}$ (Σ), $\partial'_n : E'_n \approx E'_{n-1}$ (Σ) since E, E' are elementary systems. Also $f_n : E_n \approx E'_n$ (Σ). Therefore

$$fE_{n-1} = f\partial E_n = \partial' f E_n = \partial' E'_n = E'_{n-1}$$

and

$$f_{n-1} = \partial' f_n \partial_n^{-1} : E_{n-1} \approx E'_{n-1} \ (\Sigma).$$

Therefore $f : A + C \approx A' + C'$ (Σ) induces a simple isomorphism $E \approx E'$ (Σ). Since $A + C - E = D + C$, $A' + C' - E' = D' + C'$, it follows from Lemma 1 that

$$D + C \approx D' + C' \ (\Sigma). \tag{5.9}$$

Now let $A_r \neq 0$ for some $r < p$ and let s be the least value of r with this property. Since $C'_s = 0$ and $A + C \approx A' + C'$ (Σ) it follows that s has the same property in A'. Let E now denote the direct sum of the $(s+1)$-dimensional summands in (5.8) and let D be the direct sum of the others. Let $D', E' \subset A'$ be similarly defined. Then $D_s = D'_s = 0$, by the minimal property of s. Therefore

$$(D + C)_s = (D' + C')_s = 0. \tag{5.10}$$

Since E, E' are elementary systems we have

$$\partial : E_{s+1} \approx E_s \ (\Sigma), \qquad \partial' : E'_{s+1} \approx E'_s \ (\Sigma). \tag{5.11}$$

Let $c \in C_{s+1}$, $d \in D_{s+1}$ and let $f(d + c) = e' + d' + c'$. Since $\partial'(d + c') = 0$, in consequence of (5.10), we have $\partial' e' = \partial'(e' + d' + c') = \partial' f(d + c) = f\partial(d + c) = 0$. Therefore it follows from (5.11) that $e' = 0$. Therefore $f(D + C) \subset D' + C'$. Let $h : E \to E'$ be the chain mapping induced by f. Then it follows from (5.10) that $h_s = f_s : E_s \approx E'_s$ (Σ) and also, since $he - fe \in (D' + C')_{s+1}$ if $e \in E_{s+1}$, that

[19] Cf. Theorem 1 on p. 1202 of [3].

$$\partial' h_{s+1} e = \partial' f_{s+1} e = f_s \partial e = h_s \partial e.$$

Therefore

$$h_{s+1} = \partial_{E'}^{-1} h_s \partial_E : E_{s+1} \approx E'_{s+1} \ (\Sigma),$$

where $\partial_E = \partial \mid E_{s+1}$, $\partial_{E'} = \partial' \mid E'_{s+1}$. Therefore $h : E \approx E' \ (\Sigma)$ and (5.9) again follows from Lemma 1. Lemma 3 now follows by induction on m in (5.8).

We are now approximately half way through the algebraic preliminaries. The simple homotopy equivalences will be defined as those which induce simple chain equivalences, in a sense explained in Section **10** below. But we have still to relate chain equivalences to the group T, which is defined in Section **2** above. The first step in this is to associate an element, $\tau(C) \varepsilon \mathrm{T}$, with each system, C, such that $C \equiv 0$. We shall do this by transforming C into an $(m, m+1)$-system, C^m, in which $\partial_{m+1} : C^m{}_{m+1} \approx C^m{}_m$, and defining $\tau(C) = (-1)^m \tau(\partial_{m+1})$. In Section **8** below we define a chain system called the "mapping cylinder," C^*, of a given chain equivalence $f : C \equiv C'$. This contains C as a sub-system and $C^* - C \equiv 0$. We define $\tau(f) = \tau(C^* - C)$. We shall also need to consider the effect of a Λ-automorphism, $\theta : R \approx R$, operating on T, because the chain mapping induced by a homotopy equivalence, $\phi : K \equiv K'$, is "associated" with an isomorphism $\pi_1(K) \approx \pi_1(K')$, namely the one induced by ϕ.

6. Null-equivalent systems. Let $C \equiv 0$. Then $k \simeq 1 : C \to C$, where $kC = 0$. Therefore there is a chain deformation operator, $\eta : C \to C$, such that

$$\partial\eta + \eta\partial = 1. \tag{6.1}$$

Let $\delta = \eta\partial\eta$. That is to say, $\delta = \{\delta_n\}$, where $\delta_n = \eta_n \partial_n \eta_n : C_{n-1} \to C_n$. Therefore δ is also a chain deformation operator. It follows from (6.1) that $\partial\eta\partial = (1 - \eta\partial)\partial = \partial$. Therefore $\partial\delta + \delta\partial = \partial\eta\partial\eta + \eta\partial\eta\partial = \partial\eta + \eta\partial = 1$. Also $\partial\eta\eta = (1 - \eta\partial)\eta = \eta(1 - \partial\eta) = \eta\eta\partial$. Therefore $\delta\delta = \eta\partial\eta\eta\partial\eta = \eta\eta\eta\partial\partial\eta = 0$. Thus

$$\partial\delta + \delta\partial = 1, \qquad \delta\delta = 0. \tag{6.2}$$

Let $P_1, P_2 : M \to M$ be permutations and let B be the elementary system in which

$$B_0 = 0, \quad B_i = P_i C_0, \quad B_n = 0 \qquad (i = 1, 2; n > 2)$$

and $\partial P_2 c_0 = P_1 c_0$ $(c_0 \varepsilon C_0)$. Let $C' = B + C$. Then $C'_n = C_n$ if $n \neq 1$ or 2 and

$$\begin{cases} C'_1 = P_1C_0 + C_1, & C'_2 = P_2C_0 + C_0 \\ \partial'_1(P_1c_0 + c_1) = \partial_1 c_1, & \partial'_2(P_2c_0 + c_2) = P_1c_0 + \partial_2 c_2, \end{cases}$$

where $c_i \varepsilon C_i$. Let C^* be the system which consists of the same groups, $C^*_n = C'_n$, with $\partial^*_n = \partial_n$ if $n > 2$ and

$$(6.3) \qquad \partial^*_1(P_1c_0 + c_1) = c_0, \qquad \partial^*_2(P_2c_0 + c_2) = \delta_1 c_0 + \partial_2 c_2.$$

Let $f: C' \to C^*$ be given by $f_n = 1$ if $n \neq 1$ and $f_1(P_1c_0 + c_1) = P_1\partial_1 c_1 + (\delta_1 c_0 + c_1)$. Then

$$\partial^*_1 f_1(P_1c_0 + c_1) = \partial_1 c_1 = f_0\partial'_1(P_1c_0 + c_1)$$
$$\partial^*_2 f_2(P_2c_0 + c_2) = \delta_1 c_0 + \partial_2 c_2 = P_1\partial_1\partial_2 c_2 + \delta_1 c_0 + \partial_2 c_2$$
$$= f_1(P_1c_0 + \partial_2 c_2) = f_1\partial'_2(P_2c_0 + c_2)$$

and $\partial^*_n f_n - f_{n-1}\partial'_n = \partial_n - \partial_n = 0$ if $n > 2$. Therefore f is a chain mapping. Let $g_1, h_1 : C'_1 \to C'_1$ be given by

$$g_1(P_1c_0 + c_1) = P_1(c_0 + \partial_1 c_1) + c_1$$
$$h_1(P_1c_0 + c_1) = -P_1c_0 + (\delta_1 c_0 + c_1).$$

Since $\partial_1\delta_1 = \partial_1\delta + \delta_0\partial_0 = 1$ we have

$$g_1h_1(P_1c_0 + c_1) = g_1\{- P_1c_0 + (\delta_1 c_0 + c_1)\}$$
$$= P_1(- c_0 + c_0 + \partial_1 c_1) + (\delta_1 c_0 + c_1) = f(P_1c_0 + c_1).$$

Therefore $f_1 = g_1h_1$. It follows from Theorem 2 in Section **2** that $g_1, h_1 : C'_1 \approx C'_1\ (\Sigma)$ and hence that $f: C \approx C'\ (\Sigma)$.

It follows from (6.3) that $C^* = B' + C^1$, where $B'_n = 0$ if $n > 1$,

$$B'_0 = C_0, B'_1 = P_1C_0, (\partial^* \mid B'_1) = P_1^{-1}: B'_1 \approx B'_0\ (\Sigma)$$

and

$$(6.4) \qquad C^1_0 = 0, C^1_1 = C_1, C^1_2 = P_2C_0 + C_2, C^1_n = C_n \qquad (n > 2)$$

with $\partial^1 : C^1 \to C^1$ given by $\partial^1_n = \partial_n$ if $n > 2$ and

$$(6.5) \qquad \partial^1_2(P_2c_0 + c_2) = \delta_1 c_0 + \partial_2 c_2.$$

Let $m \geq 1$ and assume that there is a system, C^m, such that $C \equiv C^m\ (\Sigma)$ and

$$C^m_0 = \cdots = C^m_{m-1} = 0, \qquad C^m_{m+p+1} = C_{m+p+1} \qquad (p > 0).$$

Then it follows from the above argument, with C^m_{m+n} playing the part of C_n, that $C^m \equiv C^{m+1}\ (\Sigma)$, where C^{m+1} satisfies the same conditions as C^m, with m

replaced by $m+1$. It follows by induction on m that there is such a system for each value of m. Moreover $C^m \equiv C \equiv 0$. Therefore equations analogous to (6.2) are satisfied in C^m.

Let $N = \dim C$ and let $m \geq N-1$. Then $C^m{}_n = 0$ unless $n = m$ or $m+1$. Therefore (6.2) reduces to

$$\text{(6.6)} \qquad \begin{cases} \partial_{m+1}\delta_{m+1} = 1 : C^m{}_m \to C^m{}_m \\ \delta_{m+1}\partial_{m+1} = 1 : C^m{}_{m+1} \to C^m{}_{m+1}. \end{cases}$$

Therefore $\partial_{m+1} : C^m{}_{m+1} \approx C^m{}_m$ and $\delta_{m+1} = \partial^{-1}{}_{m+1}$. We define the ***torsion***, $\tau(C)$, of C as

$$\text{(6.7)} \qquad \tau(C) = (-1)^m \tau(\partial_{m+1}) = (-1)^{m+1}\tau(\delta_{m+1}).$$

In (6.4) and (6.5) let C, C^1 be replaced by C^m, C^{m+1}, with $C^m{}_n = 0$ if $n > m+1$. Then (6.4), (6.5) become

$$C^{m+1}{}_{m+1} = C^m{}_{m+1}, \qquad C^m{}_{m+2} = PC^m{}_m,$$

where P is a permutation, and

$$\partial_{m+2} = \delta_{m+1}P^{-1} : C^{m+1}{}_{m+2} \to C^{m+1}{}_{m+1}.$$

Since $\tau(P) = 0$ it follows from (6.7) that $\tau(C) = (-1)^{m+1}\tau(\delta_{m+1}) = (-1)^{m+1}\tau(\partial_{m+2})$. Therefore $\tau(C)$ does not depend on the choice of $m \geq N-1$. However $\tau(C)$ appears to depend on the particular choice of δ in (6.2) and on the construction for C^m. The following theorem shows that it does not.

THEOREM 5. *$\tau(C)$ depends only on C. Also $\tau(C) = \tau(C')$, if and only if, $C \equiv C'$ (Σ), given that $C \equiv C' \equiv 0$.*

Let $C \equiv C'$ (Σ), where $C \equiv 0$, and let C^m, C'^m be any given $(m, m+1)$-systems such that $C^m \equiv C$ (Σ), $C'^m \equiv C'$ (Σ). Let $\tau(C)$ be defined by (6.7), where C^m is now this given system, and let $\tau(C')$ be similarly defined in terms of C'^m. In particular we may have $C = C'$. Therefore, when we have proved that $\tau(C) = \tau(C')$, it will follow that $\tau(C)$ depends only on C and also that $\tau(C) = \tau(C')$ if $C \equiv C'$ (Σ).

By Lemma 3 $f : B + C^m \approx B' + C'^m$ (Σ), where B, B' are collapsible, and hence elementary $(m, m+1)$-systems. It follows from Theorem 3, Corollary 2, that $B + C^m \equiv B' + C'^m \equiv 0$. Therefore it follows from relations analogous to (6.6) that

$$\partial : (B + C^m)_{m+1} \approx (B + C^m)_m$$

$$\partial' : (B' + C'^m)_{m+1} \approx (B' + C'^m)_m.$$

Moreover $\partial = f_m^{-1}\partial' f_{m+1}$ since $f\partial = \partial' f$. Since f_m, f_{m+1} are simple isomorphisms it follows that $\tau(\partial) = \tau(\partial')$. Also

$$\begin{cases} \partial(b + c) = \partial_B b + \partial_{m+1} c \\ \partial'(b' + c') = \partial_{B'} b' + \partial'_{m+1} c', \end{cases} \qquad (b \,\varepsilon\, B_{m+1}, c \,\varepsilon\, C^m{}_{m+1})$$

where $\partial_B = \partial \mid B_{m+1}$, $\partial_{B'} = \partial' \mid B'_{m+1}$. By the definition of an elementary system, $\tau(\partial_B) = \tau(\partial_{B'}) = 0$ and it follows from Theorem 2, in Section 2, that $\tau(\partial_{m+1}) = \tau(\partial) = \tau(\partial') = \tau(\partial'_{m+1})$. Therefore $\tau(C) = \tau(C')$.

Conversely let $\tau(C) = \tau(C')$ and let $C^m \equiv C\ (\Sigma)$, $C'^m \equiv C'\ (\Sigma)$, where C^m, C'^m are $(m, m+1)$-systems. Let $C^m{}_m$ be of rank p and $C'^m{}_m$ of rank p'. If $p \neq p'$, say $p < p'$, we replace C^m by $B + C^m$, where B is an elementary $(m, m+1)$-system, such that B_m is of rank $p' - p$. Therefore we assume that $p = p'$. Moreover, after applying suitable permutations to $C^m{}_m$, $C^m{}_{m+1}$ we assume that $C^m{}_m = C'^m{}_m$, $C^m{}_{m+1} = C'^m{}_{m+1}$. Then $g = \partial'_{m+1}\partial^{-1}{}_{m+1} : C^m{}_m \approx C^m{}_m$ and $(-1)^m\tau(g) = \tau(C') - \tau(C) = 0$. Therefore $g : C^m{}_m \approx C^m{}_m\ (\Sigma)$. Let $f : C^m \rightarrow C'^m$ be given by $f_m = g$, $f_{m+1} = 1$. Then $\partial'_{m+1} f_{m+1} \equiv \partial'_{m+1} = f_m \partial_{m+1}$. Therefore $f : C^m \approx C'^m\ (\Sigma)$ and the theorem is proved.

Obviously $\tau(0) = 0$. Therefore we have the corollary:

COROLLARY. $C \equiv 0\ (\Sigma)$ *if, and only if,* $\tau(C) = 0$.

Let C' be a sub-system of C, and let $C'' = C - C'$.

THEOREM 6. *If any two of C, C', C'' are chain equivalent to* 0, *so is the third and* $\tau(C) = \tau(C') + \tau(C'')$.

Let X, Y, Z denote C, C', C'' in any order and let $X \equiv Y \equiv 0$. Then $H_n(X) = 0$, $H_n(Y) = 0$ for every $n \geq 0$, according to Theorem 3, Corollary 1. Therefore it follows from the exactness of the sequence (3.4) that $H_n(Z) = 0$, for every $n \geq 0$, and hence that $Z \equiv 0$.

Let $C \equiv C' \equiv C'' \equiv 0$. Then $C \approx C' + C''\ (\Sigma)$, according to Theorem 3, Corollary 2, and Lemma 2, and it follows from Theorem 5 that $\tau(C) = \tau(C' + C'')$. For a sufficiently large value of m we have $C' \equiv A'\ (\Sigma)$, $C'' = A''\ (\Sigma)$, where A', A'' are $(m, m+1)$-systems. Therefore it follows from (5.6) that $C' + C'' \equiv A' + A''\ (\Sigma)$ and from Theorem 5 that $\tau(C) = \tau(A' + A'')$. Let $\partial' : A'_{m+1} \rightarrow A_m$, $\partial'' : A''_{m+1} \rightarrow A''_m$, be the boundary homomorphisms, which are isomorphisms since $A' \equiv A'' \equiv 0$. Then it follows from Theorems 2 and 5 that

$$\tau(A' + A'') = (-1)^m\{\tau(\partial') + \tau(\partial'')\} = \tau(A') + \tau(A'') = \tau(C') + \tau(C'')$$

and the theorem is proved.

For purposes of calculation we exhibit the structure of the system, C^m,

which is defined by reiterating the construction $C \to C^1$, leading to (6.4). To begin with we do not require $m+1 \geqq \dim C$. Let $m = 2k$ or $2k+1$ $(k \geqq 0)$ and let $D_0, D_1 \subset M$ be the basic modules

$$\left\{\begin{aligned} D_0 &= C_0 + C_2 + \cdots + C_r \\ D_1 &= C_1 + C_3 + \cdots + C_{2k+1}, \end{aligned}\right. \tag{6.8}$$

where $r = m$ if $m = 2k$, $r = m+1$ if $m = 2k+1$. Let $D = D_0 + D_1 = C_0 + C_1 + \cdots + C_{m+1}$ and let $\partial : D \to D$ be the homomorphism which is determined by $\partial : C \to C$. Let $\delta' : D \to D$ be the homomorphism which is given by

$$\begin{aligned} \delta' c &= \delta c \quad \text{if } c \in C_s \quad (s < m+1) \\ &= 0 \quad \text{if } c \in C_{m+1}. \end{aligned}$$

Then $\delta'\delta' = 0$. Also it follows from (6.2) that

$$\begin{aligned} (\partial\delta' + \delta'\partial)c &= c \quad \text{if } c \in C_s \quad (s < m+1) \\ &= \delta\partial c \quad \text{if } c \in C_{m+1}. \end{aligned} \tag{6.9}$$

Also $\partial D_i \subset D_j$, $\delta' D_i \subset D_j$ $(i \neq j; i, j = 0, 1)$. Let

$$\Delta = \partial + \delta' : (D, D_0, D_1) \to (D, D_1, D_0) \tag{6.10}$$

and let $\Delta_i : D_i \to D_j$ be the homomorphism which is induced by Δ. Then

$$\Delta_i \Delta_j d = \Delta_i \Delta d = \Delta\Delta d \qquad (d \in D_j). \tag{6.11}$$

Since $\partial\partial = \delta'\delta' = 0$ it follows from (6.9) that

$$\begin{aligned} \Delta\Delta c &= (\partial + \delta')(\partial + \delta')c \\ &= c \quad \text{if } c \in C_s \ (s < m+1); = \delta\partial c \quad \text{if } c \in C_{m+1}. \end{aligned} \tag{6.12}$$

Also $\delta'\partial C_{m+2} \subset \delta' C_{m+1} = 0$, whence $\Delta\partial C_{m+2} = 0$.

Let $i = 0$, $j = 1$ if $m = 2k$ and let $i = 1$, $j = 0$ if $m = 2k+1$. Then $C_{m+1} \subset D_j$ and it follows from (6.11) and (6.12) that

$$\begin{aligned} &\Delta_j \Delta_i = 1 \\ &\Delta_i \Delta_j c = c \quad \text{if } c \in C_s \ (s < m+1); = \delta\partial c \quad \text{if } c \in C_{m+1}. \end{aligned} \tag{6.13}$$

Let C^m, with boundary operator ∂^m, be the chain system, which is given by [20] $C^m{}_n = 0$ if $n < m$,

$$\left\{\begin{aligned} C^m{}_m &= D_i = \cdots + C_{m-2} + C_m \\ C^m{}_{m+1} &= D_j = \cdots + C_{m-1} + C_{m+1} \\ \partial^m{}_{m+1} &= \Delta_j, \end{aligned}\right. \tag{6.14}$$

[20] Cf. (5) on p. 205 of [10].

and $C^m{}_n = C_n$, $\partial^m{}_n = \partial_n$ if $n > m+1$. Since $\Delta\partial C_{m+2} = 0$ it follows that $\partial^m\partial^m = 0$. Let δ^m be the deformation operator, which is given by

$$(6.15) \qquad \begin{aligned} &\delta^m{}_{m+1} = \Delta_i \\ &\delta^m{}_{m+2}c = 0 \qquad \text{if } c \,\varepsilon\, C_s \ (s < m+1)\,; = \delta_{m+2}c \qquad \text{if } c \,\varepsilon\, C_{m+1} \end{aligned}$$

and $\delta^m{}_n = \delta_n$ if $n > m+2$. Then $\delta^m{}_{n+1}\delta^m{}_n = 0$ if $n > m+1$ and

$$\delta^m{}_{m+2}\delta^m{}_{m+1}(\cdot\cdot\cdot + c_m) = \delta_{m+2}\delta_{m+1}c_m = 0,$$

where $c_m \,\varepsilon\, C_m$. Therefore $\delta^m\delta^m = 0$. Also it follows from (6.13) that

$$\partial^m{}_{m+1}\delta^m{}_{m+1} = 1, \qquad \partial^m{}_{m+2}\delta^m{}_{m+2} + \delta^m{}_{m+1}\partial^m{}_{m+1} = 1,$$

whence $\partial^m\delta^m + \delta^m\partial^m = 1$.

Let C^{m+1} be the system, with boundary operator ∂^{m+1}, which is obtained from C^m by the construction, $C \rightarrow C^1$, leading to (6.4), with $C^m{}_{m+n}$ playing the part of C_n and with P_2C_0 replaced by C_0 and P_2 by 1 in (6.4), (6.5). Then

$$C^{m+1}{}_{m+1} = C^m{}_{m+1} = \cdot\cdot\cdot + C_{m-1} + C_{m+1}$$

$$C^{m+1}{}_{m+2} = C^m{}_m + C^m{}_{m+2} = \cdot\cdot\cdot + C_m + C_{m+2}$$

and (6.5) becomes

$$\partial^{m+1}{}_{m+2}(c^m + c_{m+2}) = \partial^m{}_{m+1}c^m + \partial_{m+2}c_{m+2} = \Delta_i c^m + \partial_{m+2}c_{m+2} = \Delta^*(c^m + c_{m+2}),$$

where $c^m \,\varepsilon\, C^m{}_m$, $c_{m+2} \,\varepsilon\, C_{m+2}$ and Δ^* is defined by (6.10), with m replaced by $m+1$. Therefore C^{m+1} is defined in the same way as C^m and we define δ^{m+1} by (6.15), with m replaced by $m+1$. Starting with $C^0 = C$, it follows by induction on m that the construction $C \rightarrow C^1$, reiterated m times, leads to C^m. Therefore $C \equiv C^m$ (Σ). We now take $m \geq \dim C - 1$, thus giving an explicit definition of ∂_{m+1}, δ_{m+1} in (6.7).

Let $R' \subset R$ be a sub-ring, which is the image of R in a homomorphism, $\phi : R \rightarrow R'$, such that $\phi r' = r'$ if $r' \,\varepsilon\, R'$. Let $1 \,\varepsilon\, R'$ and let $\Lambda' = \phi\Lambda$. It is easily verified that Λ' is a (multiplicative) group . Let $M' \subset M$ be the sub-module, which consists of all the elements $(r'_1, r'_2, \cdot\cdot\cdot)$, where $r'_i \,\varepsilon\, R'$, and let R' be such that every admissible automorphism, $M' \rightarrow M'$, is Λ-simple. That is to say, every matrix of the form (2.2), with elements in R', can be transformed into the unit matrix by a finite sequence of the transformations (2.12), with $\lambda \,\varepsilon\, \Lambda$, $r \,\varepsilon\, R$.

Let $\psi : M \rightarrow M'$ be given by $\psi(r_1, r_2, \cdot\cdot\cdot) = (\phi r_1, \phi r_2, \cdot\cdot\cdot)$. Then $\psi m_i = m_i$ and $\psi(rm) = (\phi r)\psi m$. Let $f : M \approx M$ be an admissible automorphism, which is given by $fm_i = \Sigma_j f_{ij}m_j$, $(f_{ij} \,\varepsilon\, R)$. Then $\psi f m_i = \Sigma_j(\phi f_{ij})m_j$ and it follows that $f_{ij} \,\varepsilon\, R'$ if (and only if) $\psi f = f\psi$. Therefore $f \,\varepsilon\, \Sigma$ if $\psi f = f\psi$.

Let C be a chain system, with boundary operator ∂, and let $\psi: M \to M'$ mean the same as before. Then $\psi C_r \subset C_r$, since $\psi m_i = m_i$, and $\partial\psi$, $\psi\partial$ are two families of homomorphisms $\partial\psi, \psi\partial: C_n \to C_{n-1}$.

LEMMA 4. *If $C \equiv 0$ and if $\partial\psi = \psi\partial$, then $C \equiv 0$ (Σ).*

Let $C \equiv 0$ and let $\partial\psi = \psi\partial$. Let $\eta: C \to C$ be a deformation operator, which satisfies (6.1), and let $\xi: C \to C$ be the deformation operator determined by $\xi m_i = \psi\eta m_i$, $\xi r = r\xi$. Since $\partial\psi = \psi\partial$ and $\psi m_i = m_i$ we have $(\partial\xi + \xi\partial) m_i = \psi(\partial\eta + \eta\partial) m_i = m_i$. Therefore $\partial\xi + \xi\partial = 1$. Also $\psi\xi = \xi$, since $\psi\psi = \psi$. Therefore $\xi\psi m_i = \xi m_i = \psi\xi m_i$, whence $\xi\psi = \psi\xi$. Moreover $\delta = \xi\partial\xi$ is a deformation operator such that $\delta\psi = \psi\delta$, which satisfies (6.2). It follows from (6.5) and induction on m, or from the explicit formulae (6.14), that ∂_{m+1}, in (6.6), may be constructed so as to commute with ψ. Therefore ∂_{m+1} is a simple isomorphism. Therefore $\tau(C) = 0$ and the lemma follows from the corollary to Theorem 5.

Let R be the group ring of a group Γ, let Λ consist of the elements $\pm\gamma$ ($\gamma \varepsilon \Gamma$) and let R' consist of the integral multiples of $1 \varepsilon \Gamma$. Let $\phi: R \to R'$ be given by $\phi\Gamma = 1$. Then we have the corollary:

COROLLARY. *Let $C \equiv 0$, let $(m^n_1, \cdots, m^n_{p_n})$ be the basis of C_n $(m^n_i = m_{j_i})$ and let*

$$\partial m^n_i = \Sigma_j d^n_{ij} m^{n-1}_j \qquad (n = 1, 2, \cdots),$$

where d^n_{ij} are integers. Then $C \equiv 0$ (Σ).

7. Conjugate systems. Let C be a given chain system with boundary operator ∂. Let $\theta: R \approx R$ be a given automorphism and let $s_\theta: M \to M$ mean the same as in (2.9). We shall also use s_θ to denote the semi-linear transformation $s': C_r \to C_r$, which is given by[21] $s'c = s_\theta c$ for each $c \varepsilon C_r$ $(r = 0, 1, \cdots)$. Let

$$\partial^\theta = s_\theta \partial s_\theta^{-1}: C_n \to C_{n-1} \qquad (n = 1, 2, \cdots).$$

Then $\partial^\theta\partial^\theta = s_\theta\partial\partial s_\theta^{-1} = 0$. Obviously $\partial^\theta r = r\partial^\theta$ for each operator $r \varepsilon R$. Therefore ∂^θ is a boundary operator. Let C^θ be the chain system, which consists of the modules C_n with the boundary operator ∂^θ. We shall describe C^θ as *conjugate* to C.

Let $f: C \to C'$ be a chain mapping, let $f^\theta_n = s_\theta f_n s_\theta^{-1}: C_n \to C'_n$, and let $f^\theta = \{f^\theta_n\}$. Obviously $f^\theta r = rf^\theta$ and $f^\theta\partial^\theta = s_\theta f \partial s_\theta^{-1} = s_\theta \partial f s_\theta^{-1} = \partial^\theta f^\theta$. There-

[21] $s_\theta C_r = C_r$ since $s_\theta m_i = m_i$ and C_r is a basic module in M.

fore $f^\theta: C^\theta \to C'^\theta$ is a chain mapping. On transforming the relevant equations by s_θ we see that, if $f: C \equiv C'$, then

$$f^\theta: C^\theta \equiv C'^\theta. \tag{7.1}$$

Let $C \equiv 0$ and let ∂, δ satisfy (6.2). Then ∂^θ and $\delta^\theta = s_\theta \delta s_\theta^{-1}$ obviously satisfy (6.2). It follows from (6.5) and induction on m, or from the explicit formulae (6.14), that the construction for C^m, with ∂, δ replaced by $\partial^\theta, \delta^\theta$, leads to the conjugate system $(C^m)^\theta$. Therefore

$$\tau(C^\theta) = (-1)^m \tau(\partial^\theta{}_{m+1}). \tag{7.2}$$

Let $\theta\Lambda = \Lambda$. Then it follows from (7.2) that

$$\tau(C^\theta) = \theta\tau(C), \tag{7.3}$$

where $\theta: \mathrm{T} \to \mathrm{T}$ is defined by (2.10).

8. Mapping cylinders. Let C, C' be disjoint chain systems, with boundary operators ∂, ∂', and let $f: C \to C'$ be a chain mapping. Let αC_{n-1} be the image of C_{n-1} in a simple isomorphism $\alpha: C_{n-} \approx \alpha C_{n-1}$, which is induced by a permutation $P_{n-1}: M \to M$. Then αC_{n-1} is a basic module. Let $C^*{}_n$ be the direct sum $C^*{}_n = C'_n + C_n + \alpha C_{n-1}$. Let $\partial: C^*{}_n \to C^*{}_{n-1}$ be defined by

$$\left\{\begin{array}{lll} \text{a)} & \partial^* c = \partial c, \quad \partial^* c' = \partial' c' & (c \in C_n,\ c' \in C'_n) \\ \text{b)} & \partial^* \alpha c = (f - 1 - \alpha\partial) c & (c \in C_{n-1}). \end{array}\right. \tag{8.1}$$

We shall write (8.1b) as $\partial^*\alpha = f - 1 - \alpha\partial$, using 1, f as abbreviations for i, $i'f$, where $i: C_n \to C^*{}_n$, $i': C'_n \to C^*{}_n$ are the identities. Obviously $\partial^*\partial^*(C'_n + C_n) = 0$ and $\partial^*\partial^*\alpha = \partial' f - \partial - \partial^*\alpha\partial = (f-1)\partial - (f - 1 - \alpha\partial)\partial = 0$. Therefore $C^* = \{C^*{}_n\}$, with ∂^* as boundary operator, is a chain system. We shall call it the *mapping cylinder* of f. Clearly C, C' are subsystems of C.

LEMMA 5. *$C^* - C'$ is collapsible.*

Let $C'' = C^* - C'$. Then $C''_n = C_n + \alpha C_{n-1}$ and

$$\partial''(c_2 + \alpha c_1) = \partial c_2 - (1 + \alpha\partial) c_1 = (\partial c_2 - c_1) - \alpha\partial c_1, \tag{8.2}$$

where $c_2 \in C_n$, $c_1 \in C_{n-1}$. Let $\partial^0: C'' \to C''$ be given by

$$\partial^0 c = 0, \qquad \partial^0 \alpha c = c.$$

Then $\partial^0\partial^0 = 0$ and it follows that $\{C''_n\}$, with ∂^0 as boundary operator, is a chain system C^0, which is obviously collapsible. Let $g: C'' \to C^0$ be given by

$g(c_2 + \alpha c_1) = c_2 + \alpha(\partial c_2 - c_1)$. Then $g\partial''(c_2 + \alpha c_1) = g(\partial c_2 - c_1 - \alpha\partial c_1)$ $= \partial c_2 - c_1 + \alpha(-\partial c_1 + \partial c_1) = \partial^0 g(c_2 + \alpha c_1)$. Therefore it follows from Lemma 1, Section **3**, that $g: C'' \approx C^0$ (Σ), and Lemma 5 is proved.

It follows from Lemma 5 and Theorem 4 that $C^* \equiv C'$ (Σ), rel. C'. Therefore C' is a D. R. of C^* and any retraction $k': C^* \to C'$ is a simple equivalence. Let k' be given by $k'c = fc$, $k'c' = c'$, $k'\alpha c = 0$ $(c \in C, c' \in C')$. Then it follows from (8.1) that $k'\partial^* = \partial' k'$. Therefore k' is a chain mapping, which is a retraction, since $k'c' = c'$. Also $k'c = fc = fic$. Therefore we have the corollary.

COROLLARY. *$k': C^* \equiv C'$ (Σ), rel. C', and $f = k'i$.*

Let $C(f) = C^* - C$. Then $C_n(f) = C'_n + \alpha C_{n-1}$, and it follows from (8.1) that the boundary operator, $\partial^f: C(f) \to C(f)$, is given by

$$\partial^f c' = \partial' c', \qquad \partial^f \alpha = f - \alpha\partial. \tag{8.3}$$

LEMMA 6. *$C(f) \equiv 0$ if, and only if, $f: C \equiv C'$.*

Let $C(f) \equiv 0$. Then C is a D. R. of C^*, according to Theorem 3, Corollary 2. Therefore $i: C \equiv C^*$ and, by the corollary to Lemma 5, $k': C^* \equiv C'$. Therefore $f: C \equiv C'$.

Conversely, let $f: C \equiv C'$. Then

$$f_* = k'_* i_*: H_n(C) \approx H_n(C') \qquad (n = 0, 1, \cdots)$$

and it follows that $i_*: H_n(C) \approx H_n(C^*)$. Therefore it follows from the exactness of the sequence

$$H_n(C) \to H_n(C^*) \to H_n\{C(f)\} \to H_{n-1}(C) \to H_{n-1}(C^*)$$

that $H_n\{C(f)\} = 0$ for every[22] $n \geq 0$. Therefore $C(f) \equiv 0$ by Theorem 3, Corollary 1.

LEMMA 7. *If $f \simeq g: C \to C'$, then $C(f) \approx C(g)$ (Σ).*

Let $g - f = \partial\eta + \eta\partial$, where $\eta: C \to C'$ is a deformation operator, and let βC_n and $\beta: C_n \approx \beta C_n$ be the analogues, in $C(g)$, of αC_n and α. Let $h: C(f) \to C(g)$ be given by $h(c' + \alpha c) = (c' - \eta c) + \beta c$, which we write as $hc' = c'$, $h\alpha = \beta - \eta$. Then $\partial h = h\partial$ in C' and since $hf = f$ we have[23]

$$\partial h\alpha = \partial(\beta - \eta) = g - \beta\partial - (g - f - \eta\partial) = f - (\beta - \eta)\partial = hf - h\alpha\partial = h\partial\alpha.$$

Therefore it follows from Lemma 1 that $h: C(f) \approx C(g)$ (Σ) and the lemma is proved.

---

[22] Cf. Section **3** in [5].

[23] We now use ∂ to denote the boundary operator in all our systems.

Let $f: C \equiv C'$. Then it follows from Lemma 6 that $C(f) \equiv 0$. We define $\tau(f) = \tau\{C(f)\}$ and call $\tau(f)$ the *torsion* of f. It follows from Lemma 7 that $\tau(f)$ depends only on the chain-homotopy class containing f and we shall also call it the *torsion* of this class.

Let $f: C \to C'$, $f': C' \to C''$ be any chain mappings.

LEMMA 8. *There is a chain system D, containing $C(f)$ as a sub-system, such that*[24] $D - C(f) = C(f')$ *and* $D \equiv C(f'f)$ (Σ).

Let $\alpha' C'_n \subset C(f')$ and $\alpha'' C_n \subset C(f'f)$ be the analogues of αC_n. Let C'^* be the mapping cylinder of f' and let D be the direct sum $D = C'^* + C(f)$ with the "united sub-system" C'. That is to say $D_n = C''_n + C'_n + \alpha' C'_{n-1} + \alpha C_{n-1}$ and $\partial: D \to D$ is determined by $\partial: C'^* \to C'^*$ and $\partial: C(f) \to C(f)$, which coincide in C'. Thus $\partial\alpha' = f' - 1 - \alpha'\partial$, $\partial\alpha = f - \alpha\partial$. Moreover $C(f)$ is a sub-system of D and obviously $D - C(f) = C(f')$.

Let D' be the direct sum $D' = C'^* + C(f'f)$, with the united sub-system C''. Then $D'_n = C''_n + C'_n + \alpha' C'_{n-1} + \alpha'' C_{n-1}$ and $\partial\alpha'' = f'f - \alpha''\partial$.

Let $g: D \to D'$ be given by

$$g(c'^* + \alpha c) = (c'^* - \alpha' f c) + \alpha'' c \qquad (c \,\varepsilon\, C, c'^* \,\varepsilon\, C'^*),$$

which we write as $gc'^* = c'^*$, $g\alpha = \alpha'' - \alpha' f$. Then $\partial g = g\partial$ in C'^*. Since $gfc = fc$ we have

$$\begin{aligned} \partial g\alpha = \partial\alpha'' - \partial\alpha' f = f'f - \alpha''\partial - (f' - 1 - \alpha'\partial)f &= f - \alpha''\partial + \alpha'\partial f \\ &= f - (\alpha'' - \alpha' f)\partial = g(f - \alpha\partial) = g\partial\alpha. \end{aligned}$$

Therefore $g: D \approx D'$ (Σ). Clearly $D' - C(f'f) = C'^* - C''$ and it follows from Lemma 5 and Theorem 4(b) that $C(f'f) \equiv D' \approx D$ (Σ). This completes the proof.

Let $f: C \equiv C'$, $f': C' \equiv C''$.

THEOREM 7. $\tau(f'f) = \tau(f') + \tau(f)$.

Let D mean the same as in Lemma 8. Then it follows from Lemma 8 and Theorems 5, 6 that $\tau\{C(f'f)\} = \tau(D) = \tau\{C(f')\} + \tau\{C(f)\}$, and the theorem is proved.

LEMMA 9. *If* $f: C \approx C'$ (Σ) *then* $\tau(f) = 0$.

Let $f: C \approx C'$ (Σ) and let C^* be the mapping cylinder of f. Let $g: C(f) \to C^* - C'$ be given by $g(c' + \alpha c) = f^{-1}c' - \alpha c$. Then $\partial g c' = g\partial c'$ and it follows from (8.2) that

[24] I. e. $D - C(f) = C(f')$ if the permutations $\alpha': C'_n \to \alpha' C'_n \subset C(f')$ are suitably chosen.

$$\partial g\alpha = 1 + \alpha\partial = f^{-1}f + \alpha\partial = g(f - \alpha\partial) = g\partial\alpha.$$

Therefore $g: C(f) \approx C^* - C'$ (Σ) and the lemma follows from Lemma 5 and the corollary to Theorem 5.

Let $f: C \equiv C'$ and let us discard the (implicit) condition that $C \cap C' = 0$. Let $h: A \approx C$ (Σ), where A is a chain system such that $A \cap C' = 0$. Then $fh: A \equiv C'$ and we define $\tau(f)$ by $\tau(f) = \tau(fh)$. Let (h', A) be any other pair such that $h': A \approx C$ (Σ), $A' \cap C' = 0$. Let $h'': A'' \approx C$ (Σ), where A'' is disjoint from A, A', C'. Then $fh: A \equiv C'$, $h^{-1}h'': A'' \approx A$ (Σ), $fh'': A'' \equiv C'$. Therefore it follows from Theorem 7 and Lemma 9 that $\tau(fh'') = \tau(fh) + \tau(h^{-1}h'') = \tau(fh)$. Similarly $\tau(fh'') = \tau(fh')$. Therefore $\tau(f)$ is independent of the choice of h, A.

Let $f \simeq g: C \equiv C'$, where $C \cap C' \neq 0$. Then $fh \cong gh: A \equiv C'$, and in consequence of Lemma 7 we have:

THEOREM 8. *If $f \simeq g: C \equiv C'$, then $\tau(f) = \tau(g)$.*

Let $f: C \equiv C'$, $f': C' \equiv C''$. Let $h: A \approx C$ (Σ), $h': A' \approx C'$ (Σ), where A. A' are disjoint from C', C'' and from each other. Then $f'h': A' \equiv C''$, $h'^{-1}fh: A \equiv A'$, $f'fh: A \equiv C''$, and $h'^{-1}: C' \approx A'$ (Σ), $fh: A \equiv C'$. Therefore it follows from Theorem 7, and Lemma 9 that

$$\begin{aligned}\tau(f'f) = \tau(f'fh) = \tau(f'h' \cdot h'^{-1}fh) &= \tau(f'h') + \tau(h'^{-1}fh) \\ &= \tau(f'h') - \tau(h') + \tau(fh) = \tau(f') + \tau(f).\end{aligned}$$

Therefore Theorem 7 is valid, even when C, C', C'' are not disjoint from each other.

Similarly Lemma 9 is valid, even if $C \cap C' \neq 0$.

THEOREM 9. *Given $g: C \equiv C'$, then $\tau(g) = 0$ if, and only if, $g: C \equiv C'$ (Σ).*

It follows from Theorem 7 and Lemma 9 that we may assume $C \cap C' = 0$. This being so, let $\tau(g) = 0$ and let C^* be the mapping cylinder of g. Since $\tau\{C(g)\} = \tau(g) = 0$ it follows from Theorem 5 that $C(g) \equiv 0$ (Σ). Therefore it follows from Theorem 4(b) that $C^* \equiv C$ (Σ), rel. C, whence $i: C \equiv C^*$ (Σ). Therefore it follows from Corollary to Lemma 5 that $g: C \equiv C'$ (Σ).

Conversely, let $g: C \equiv C'$ (Σ). Then $f: B + C \approx B' + C'$ (Σ), $g \simeq k'fi$, where B, B' are collapsible systems, $i: C \to B + C$ is the identity and $k': B' + C' \to C'$ is a retraction. Assume that $\tau(i) = \tau(i') = 0$, where $i': C' \to B' + C'$ is the identity. Then $\tau(k') = 0$, since $\tau(k') + \tau(i') = \tau(k'i') = \tau(1) = 0$. Also $\tau(f) = 0$, according to Lemma 9, and it follows from Theorems 7, 8 that $\tau(g) = 0$.

Let $f: C \equiv C'$. Then it follows from Lemma 6 that $C(f) \equiv 0$. We define $\tau(f) = \tau\{C(f)\}$ and call $\tau(f)$ the *torsion* of f. It follows from Lemma 7 that $\tau(f)$ depends only on the chain-homotopy class containing f and we shall also call it the *torsion* of this class.

Let $f: C \rightarrow C'$, $f': C' \rightarrow C''$ be any chain mappings.

LEMMA 8. *There is a chain system D, containing $C(f)$ as a sub-system, such that*[24] $D - C(f) = C(f')$ *and* $D \equiv C(f'f)$ (Σ).

Let $\alpha' C'_n \subset C(f')$ and $\alpha'' C_n \subset C(f'f)$ be the analogues of αC_n. Let C'^* be the mapping cylinder of f' and let D be the direct sum $D = C'^* + C(f)$ with the "united sub-system" C'. That is to say $D_n = C''_n + C'_n + \alpha' C'_{n-1} + \alpha C_{n-1}$ and $\partial: D \rightarrow D$ is determined by $\partial: C'^* \rightarrow C'^*$ and $\partial: C(f) \rightarrow C(f)$, which coincide in C'. Thus $\partial\alpha' = f' - 1 - \alpha'\partial$, $\partial\alpha = f - \alpha\partial$. Moreover $C(f)$ is a sub-system of D and obviously $D - C(f) = C(f')$.

Let D' be the direct sum $D' = C'^* + C(f'f)$, with the united sub-system C''. Then $D'_n = C''_n + C'_n + \alpha' C'_{n-1} + \alpha'' C_{n-1}$ and $\partial\alpha'' = f'f - \alpha''\partial$.

Let $g: D \rightarrow D'$ be given by

$$g(c'^* + \alpha c) = (c'^* - \alpha' f c) + \alpha'' c \qquad (c \in C, c'^* \in C'^*),$$

which we write as $gc'^* = c'^*$, $g\alpha = \alpha'' - \alpha' f$. Then $\partial g = g\partial$ in C'^*. Since $gfc = fc$ we have

$$\begin{aligned} \partial g\alpha = \partial\alpha'' - \partial\alpha' f = f'f - \alpha''\partial - (f' - 1 - \alpha'\partial)f = f - \alpha''\partial + \alpha'\partial f \\ = f - (\alpha'' - \alpha' f)\partial = g(f - \alpha\partial) = g\partial\alpha. \end{aligned}$$

Therefore $g: D \approx D'$ (Σ). Clearly $D' - C(f'f) = C'^* - C''$ and it follows from Lemma 5 and Theorem 4(b) that $C(f'f) \equiv D' \approx D$ (Σ). This completes the proof.

Let $f: C \equiv C'$, $f': C' \equiv C''$.

THEOREM 7. $\tau(f'f) = \tau(f') + \tau(f)$.

Let D mean the same as in Lemma 8. Then it follows from Lemma 8 and Theorems 5, 6 that $\tau\{C(f'f)\} = \tau(D) = \tau\{C(f')\} + \tau\{C(f)\}$, and the theorem is proved.

LEMMA 9. *If* $f: C \approx C'$ (Σ) *then* $\tau(f) = 0$.

Let $f: C \approx C'$ (Σ) and let C^* be the mapping cylinder of f. Let $g: C(f) \rightarrow C^* - C'$ be given by $g(c' + \alpha c) = f^{-1}c' - \alpha c$. Then $\partial g c' = g\partial c'$ and it follows from (8.2) that

[24] I. e. $D - C(f) = C(f')$ if the permutations $\alpha': C'_n \rightarrow \alpha' C'_n \subset C(f')$ are suitably chosen.

$$\partial g\alpha = 1 + \alpha\partial = f^{-1}f + \alpha\partial = g(f - \alpha\partial) = g\partial\alpha.$$

Therefore $g: C(f) \approx C^* - C'$ (Σ) and the lemma follows from Lemma 5 and the corollary to Theorem 5.

Let $f: C \equiv C'$ and let us discard the (implicit) condition that $C \cap C' = 0$. Let $h: A \approx C$ (Σ), where A is a chain system such that $A \cap C' = 0$. Then $fh: A \equiv C'$ and we define $\tau(f)$ by $\tau(f) = \tau(fh)$. Let (h', A) be any other pair such that $h': A \approx C$ (Σ), $A' \cap C' = 0$. Let $h'': A'' \approx C$ (Σ), where A'' is disjoint from A, A', C'. Then $fh: A \equiv C'$, $h^{-1}h'': A'' \approx A$ (Σ), $fh'': A'' \equiv C'$. Therefore it follows from Theorem 7 and Lemma 9 that $\tau(fh'') = \tau(fh) + \tau(h^{-1}h'') = \tau(fh)$. Similarly $\tau(fh'') = \tau(fh')$. Therefore $\tau(f)$ is independent of the choice of h, A.

Let $f \simeq g: C \equiv C'$, where $C \cap C' \neq 0$. Then $fh \cong gh: A \equiv C'$, and in consequence of Lemma 7 we have:

THEOREM 8. *If $f \simeq g: C \equiv C'$, then $\tau(f) = \tau(g)$.*

Let $f: C \equiv C'$, $f': C' \equiv C''$. Let $h: A \approx C$ (Σ), $h': A' \approx C'$ (Σ), where A. A' are disjoint from C', C'' and from each other. Then $f'h': A' \equiv C''$, $h'^{-1}fh: A \equiv A'$, $f'fh: A \equiv C''$, and $h'^{-1}: C' \approx A'$ (Σ), $fh: A \equiv C'$. Therefore it follows from Theorem 7, and Lemma 9 that

$$\begin{aligned}\tau(f'f) = \tau(f'fh) = \tau(f'h' \cdot h'^{-1}fh) &= \tau(f'h') + \tau(h'^{-1}fh) \\ &= \tau(f'h') - \tau(h') + \tau(fh) = \tau(f') + \tau(f).\end{aligned}$$

Therefore Theorem 7 is valid, even when C, C', C'' are not disjoint from each other.

Similarly Lemma 9 is valid, even if $C \cap C' \neq 0$.

THEOREM 9. *Given $g: C \equiv C'$, then $\tau(g) = 0$ if, and only if, $g: C \equiv C'$ (Σ).*

It follows from Theorem 7 and Lemma 9 that we may assume $C \cap C' = 0$. This being so, let $\tau(g) = 0$ and let C^* be the mapping cylinder of g. Since $\tau\{C(g)\} = \tau(g) = 0$ it follows from Theorem 5 that $C(g) \equiv 0$ (Σ). Therefore it follows from Theorem 4(b) that $C^* \equiv C$ (Σ), rel. C, whence $i: C \equiv C^*$ (Σ). Therefore it follows from Corollary to Lemma 5 that $g: C \equiv C'$ (Σ).

Conversely, let $g: C \equiv C'$ (Σ). Then $f: B + C \approx B' + C'$ (Σ), $g \simeq k'fi$, where B, B' are collapsible systems, $i: C \to B + C$ is the identity and $k': B' + C' \to C'$ is a retraction. Assume that $\tau(i) = \tau(i') = 0$, where $i': C' \to B' + C'$ is the identity. Then $\tau(k') = 0$, since $\tau(k') + \tau(i') = \tau(k'i') = \tau(1) = 0$. Also $\tau(f) = 0$, according to Lemma 9, and it follows from Theorems 7, 8 that $\tau(g) = 0$.

It remains to prove that $\tau(i) = \tau(i') = 0$. Let $h: A \approx C$ (Σ), where $A \cap (B + C) = 0$. Since $ihA = C$ it follows from (8.3) that $C(ih) = B + C(h)$, and from Theorems 5, 6 and Lemma 6 that $\tau\{C(ih)\} = \tau(B) + \tau\{C(h)\} = 0$. Therefore $\tau(i) = \tau(ih) = 0$ and similarly $\tau(i') = 0$. This completes the proof.

Let A, A' be sub-systems of C, C' and let $h: C \rightarrow C'$ be a chain mapping such that $hA \subset A'$. Let $B = C - A$, $B' = C' - A'$ and let $f: A \rightarrow A'$, $g: B \rightarrow B'$, be the chain mappings induced by h.

THEOREM 10. *If any two of f, g, h are chain equivalences so is the third, and $\tau(h) = \tau(f) + \tau(g)$.*

Assuming that $C \cap C' = 0$ we have

$$C_n(h) = C'_n + \alpha C_{n-1} = A'_n + B'_n + \alpha A_{n-1} + \alpha B_{n-1} = C_n(f) + C_n(g).$$

Let $D = C(h) - C(f)$. Then $D_n = C_n(g)$ and I say that $D = C(g)$. For let ∂_X denote the boundary operator in X, where X stands for any of the systems $C, C', \cdots$. Since $A' \subset C(f)$ we have [25]

$$\partial_D b' \equiv \partial_{C(h)} b' = \partial_{C'} b' \equiv \partial_{B'} b' = \partial_{C(g)} b', \qquad \text{mod. } C(f),$$

where $b' \varepsilon B'$. Since $\alpha A \subset C(f)$ we have $\alpha c_1 \equiv \alpha c_2$, mod. $C(f)$, if $c_1 \equiv c_2$, mod. A, where $c_1, c_2 \subset C$. Therefore

$$\partial_D \alpha b \equiv \partial_{C(h)} \alpha b = hb - \alpha \partial_C b \equiv gb - \alpha \partial_B b = \partial_{C(g)} b, \qquad \text{mod. } C(f),$$

where $b \varepsilon$ B. Therefore $\partial_D d \equiv \partial_{C(g)} d$, mod. $C(f)$, for any $d \varepsilon D$. But $D \cap C(f) = 0$. Therefore $\partial_D = \partial_{C(g)}$, whence $D = C(g)$. The theorem now follows from Lemma 6 and Theorem 6.

In consequence of Theorem 6 we have:

COROLLARY. *If any two of f, g, h are simple equivalences, so is the third.*

Using the same notation as in Theorem 10, let $A' = A = C \cap C'$ and let $h: C \rightarrow C'$ be rel. A. Then we define the *mapping cylinder*, C^*, of h in the same way as when $A = 0$, except that $C^*_n = C'_n + B_n + \alpha B_{n-1} = B'_n + C_n + \alpha B_{n-1}$ where $\alpha: B_n \approx \alpha B_n$ (Σ) is induced by a permutation, $P_n: M \rightarrow M$, and ∂^* is given by (8.1), with $c \varepsilon B$ and $\partial = \partial_B$ in (8.1b). Obviously $C^* - C = C(g)$. Therefore it follows from Theorem 10, with $f = 1$, and from Lemma 6, that $h: C \equiv C'$ if, and only if, $C^* - C \equiv 0$. It

[25] If Y is a sub-system of X, then $x \equiv x'$, mod $Y (x, x' \subset X)$ means that $x - x' \varepsilon Y$. Clearly $\partial_z x \equiv \partial_x x$, mod Y, where $Z = X - Y$.

follows from the corollary to Theorem 10 and Theorem 9 that h is a simple equivalence if, and only if,

$$C^* - C \equiv 0 \ (\Sigma). \tag{8.4}$$

Let $f: C \equiv C'$ and let $\theta: R \approx R$ be any Λ-automorphism. Then $f^\theta: C^\theta \equiv C'^\theta$, according to (7.1). Since α, in (8.1), is the isomorphism induced by a permutation it follows that $\alpha s_\theta = s_\theta \alpha$. Therefore it follows from (8.3) that $C(f^\theta) = C(f)^\theta$ and from (7.3) that

$$\tau(f^\theta) = \theta\tau(f). \tag{8.5}$$

Let $f: C \equiv C'$. In order to calculate $\tau(f)$ we need to know a deformation operator, $\xi: C(f) \to C(f)$, such that $\partial\xi + \xi\partial = 1$. Let f and $f': C' \to C$ be related by $f'f - 1 = \partial\eta + \eta\partial$, $ff' - 1 = \partial\eta' + \eta'\partial$, where $\eta: C \to C$, $\eta': C' \to C'$ are deformation operators. Let

$$\mu = f\eta - \eta' f: C \to C'. \tag{8.6}$$

Then

$$\begin{aligned} \partial\mu + \mu\partial &= \partial f\eta - \partial\eta' f + f\eta\partial - \eta' f\partial = f(\partial\eta + \eta\partial) - (\partial\eta' + \eta'\partial)f \\ &= f(f'f - 1) - (ff' - 1)f = 0. \end{aligned}$$

Hence it follows by a straightforward calculation that $\partial\xi + \xi\partial = 1$, where $\xi: C(f) \to C(f)$ is given by

$$\begin{cases} \xi\alpha = \alpha\eta - \mu\eta \\ \xi \mid C' = \alpha f' - \mu f' - \eta'. \end{cases} \tag{8.7}$$

9. The groupoid $\mathfrak{G}$. Let R be the integral group ring of a group Γ. We need to consider chain mappings which do not necessarily commute with the operators $r \in R$. By a *chain mapping, $f: C \to C'$, associated with an automorphism, $\theta: R \approx R$,* we shall mean a family of homomorphisms, $f_n: C_n \to C'_n$, such that $f\partial = \partial f$ and $fr = (\theta r)f$. We now insist that $C_0 \neq 0$ and that, if m_i is a basis element of C_0, then $f_0 m_i$ shall be a basis element of C'_0. This ensures that f is associated with only one automorphism θ (unlike $C \to 0$, for example). If $f': C' \to C''$ is associated with $\theta': R \approx R$, then $f'f: C \to C''$ is obviously associated with $\theta'\theta$. Let $x \in R$ be any regular element. Then $\partial x = x\partial$ and $x(rc) = (xrx^{-1})xc$. Therefore $x: C \to C$, given by $c \to xc$, is a chain mapping associated with the inner automorphism θ_x. We shall confine ourselves to chain mappings associated with those automorphisms of R, which are determined by automorphisms of Γ, and we shall

use the same symbol to denote $\theta:\Gamma\approx\Gamma$ and the corresponding automorphism of R.

We define chain homotopy and chain equivalence as in CH(II), with Γ playing the part of $\bar{\rho}_1$. Thus $f\simeq g:C\to C'$ means that

$$(9.1)\qquad \gamma g-f=\partial\eta+\eta\partial,$$

where $\gamma\,\varepsilon\,\Gamma$ and $\eta:C\to C'$ is a chain deformation operator associated with the same automorphism, θ, as f. As in CH(II) it follows that g is associated with $\theta_{\gamma}{}^{-1}\theta$. We shall write $f\cong g$ if, and only if, f, g are related by (9.1), with $\gamma=1$. As in the ordinary theory of homotopy or chain homotopy, $f\simeq g$ implies $fh\simeq gh$, $h'f\simeq h'g$, where h, h' are any chain mappings of the form $h:C^0\to C$, $h':C'\to C''$.

We say that $f:C\to C'$ is a *chain equivalence* and write $f:C\equiv C'$, if, and only if there is a chain mapping, $g:C'\to C$, such that $gf\simeq 1$, $fg\simeq 1$. Let $\gamma,\gamma'\subset\Gamma$ be such that $\gamma gf\cong 1$, $\gamma' fg\cong 1$. Then $f'f\cong 1$, where $f'=\gamma g$. On transforming $\gamma' fg\cong 1$ by γ' we have $fg\gamma'\cong 1$. Therefore $ff'\cong ff'fg\gamma'\cong fg\gamma'\cong 1$.

Let $f:C\equiv C'$, where f is associated with $1:R\approx R$. That is to say, $fr=rf$. Let $f':C'\to C$ be such that $f'f\cong 1$, $ff'\cong 1$. Since $f'f$ is associated with only one $\theta:R\approx R$ and since $fr=rf$ it follows that $f'r=rf'$. Therefore $f:C\equiv C'$, in the sense of Lemma 6.

Let $f:C\equiv C'$ be associated with $\theta:R\approx R$ and let $\theta_1,\theta_2:R\approx R$ be given. Using the same notation as in Section 7 we have

$$\partial'^{\theta_1}s_{\theta_1}fs_{\theta_2}{}^{-1}=s_{\theta_1}\partial'fs_{\theta_2}{}^{-1}=s_{\theta_1}f\partial s_{\theta_2}{}^{-1}=s_{\theta_1}fs_{\theta_2}{}^{-1}\partial^{\theta_2}.$$

Therefore

$$(9.2)\qquad s_{\theta_1}fs_{\theta_2}{}^{-1}:C^{\theta_2}\to C'^{\theta_1}$$

is a chain mapping, which is obviously associated with $\theta_1\theta\theta_2{}^{-1}$. Let $f':C'\to C$ be such that $f'f\cong 1$, $ff'\cong 1$. On transforming $f'f\cong 1$ by s_θ we have $s_\theta f'fs_\theta{}^{-1}\cong 1:C^\theta\to C^\theta$. Also $fs_\theta{}^{-1}\cdot s_\theta f'\cong 1:C'\to C'$. Therefore it follows from (9.2), with $\theta_1=1$, $\theta_2=\theta$, that $fs_\theta{}^{-1}:C^\theta\equiv C'$. Moreover $(fs_\theta{}^{-1})r=r(fs_\theta{}^{-1})$. We define $\tau(f)$ by

$$(9.3)\qquad \tau(f)=\tau(fs_\theta{}^{-1})$$

and call it the *torsion* of f. Let $\theta_1,\theta_2:R\approx R$ be arbitrary. Then $s_{\theta_1}fs_{\theta_2}{}^{-1}$ is associated with $\theta_1\theta\theta_2{}^{-1}$ and it follows from (9.3) and (8.5) that

$$(9.4)\qquad \tau(s_{\theta_1}fs_{\theta_2}{}^{-1})=\tau(s_{\theta_1}fs_{\theta_2}{}^{-1}\cdot s_{\theta_2}s_\theta{}^{-1}s_{\theta_1}{}^{-1})=\tau(s_{\theta_1}\cdot fs_\theta{}^{-1}\cdot s_{\theta_1}{}^{-1})=\theta_1\tau(f).$$

Let $f' : C' \equiv C''$ and let f' be associated with $\theta' : R \approx R$. Then $f'f : C \equiv C''$ is associated with $\theta'\theta$, and it follows from Theorem 7 and (9.4) that

$$\tau(f'f) = \tau(f'fs_\theta^{-1}s_{\theta'}^{-1}) = \tau(f's_{\theta'}^{-1} \cdot s_{\theta'}fs_\theta^{-1}\, s_{\theta'}^{-1}) \\ = \tau(f') + \tau(s_{\theta'} \cdot fs_\theta^{-1} \cdot s_{\theta'}^{-1}) = \tau(f') + \theta'\tau(f). \tag{9.5}$$

Let $f \cong g : C \equiv C'$ and let f, g be related by (9.1). Then $\gamma g, f$ and η are all associated with the same automorphism, θ, and

$$\gamma g s_\theta^{-1} - f s_\theta^{-1} = \partial'\eta s_\theta^{-1} + \eta s_\theta^{-1} s_\theta \partial s_\theta^{-1} = \partial'\xi + \xi\partial^\theta,$$

where $\xi = \eta s_\theta^{-1}$. Therefore $fs_\theta^{-1} \cong \gamma g s_\theta^{-1}$ and it follows from Theorem 8 that $\tau(f) = \tau(\gamma g)$. Also it follows from (9.5) and (2.11) that $\tau(\gamma g) = \tau(\gamma) + \theta_\gamma\tau(g) = \tau(\gamma) + \tau(g)$, where $\gamma : C' \to C'$ is the chain mapping $c' \to \gamma c'$, which is obviously a chain equivalence. Since $\gamma s_{\theta\gamma}{}^{-1}m_i = \gamma m_i$, where m_i is any basis element of C'_n, it follows from (2.7) that $\gamma s_{\theta\gamma}{}^{-1} : C'^{\theta\gamma} \approx C'\ (\Sigma)$, whence $\tau(\gamma) = 0$. Therefore $f \cong g$ implies

$$\tau(f) = \tau(g). \tag{9.6}$$

Let $\mathfrak{G}$ be the totality of chain homotopy classes of equivalences between all the chain systems, which are equivalent to a given one. Let $f : C \equiv C'$, $f' : C' \equiv C''$ be such equivalences and $\bar{f}$, $\bar{f}'$ the corresponding chain homotopy classes. We define $\bar{f}'\bar{f} = \overline{f'f}$. It may be verified that, when multiplication is thus defined, $\mathfrak{G}$ is a groupoid. Let f be associated with θ. Then we define $f : \mathrm{T} \to \mathrm{T}$ by $f\tau = \theta\tau$. Let $f \cong g$. Then g is associated with $\theta_\gamma{}^{-1}\theta$, for some $\gamma \varepsilon \Gamma$, and it follows from (2.11) that $g\tau = \theta_\gamma{}^{-1}\theta\tau = \theta\tau = f\tau$. Therefore a single-valued map $\bar{f}$: $\mathrm{T} \to \mathrm{T}$ is defined by $\bar{f}\tau = f\tau$. Obviously $1 \cdot \tau = \tau$ if 1 is any identical map, $C \to C$. Since $f'f$, if it exists, is associated with $\theta'\theta$, where f, f' are associated with θ, θ', it follows that $\bar{f}'(\bar{f}\tau) = (\bar{f}'\bar{f})\tau$. Therefore we say that T admits $\mathfrak{G}$ as a *groupoid of operators.*

It follows from (9.6) that a single-valued map $\tau : \mathfrak{G} \to \mathrm{T}$ is defined by $\tau(\bar{f}) = \tau(f)$ and from (9.5) that

$$\tau(\bar{g}\bar{f}) = \tau(\bar{g}) + \bar{g}\tau(\bar{f}). \tag{9.7}$$

Therefore τ is what, by a natural extension of the language of group theory, we call a *crossed homomorphism* of $\mathfrak{G}$ into T. We call $\tau(\bar{f})$ the *torsion* of $\bar{f}$. Given C it is easy to construct a chain system $C' \equiv C$ and an equivalence, $f : C \equiv C'$, such that $\tau(f)$ is a given element $\tau_0 \varepsilon \mathrm{T}$. For example, let $d : A \approx B$ be an isomorphism such that $\tau(d) = \tau_0$, where A, B are basic modules, which are disjoint from C and from each other Let $m > \dim C$ and let C' be the system which consists of C, with its own boundary operator, and

$C'_m = B$, $C'_{m+1} = A$, with $\partial'_{m+1} = d$. Then it follows from Theorem 10 that $\tau(i) = \tau_0$, where i is the identical map $C \to C'$.

10. Homotopy types of complexes. Let K be a given complex and let a 0-cell $e^0 \varepsilon K^0$ be taken as base-point for $\pi_1(K)$. Let $\tilde{K}$ be the universal covering complex of K, in which the points are classes of paths joining e^0 to points in K. Let[26] $C(\tilde{K})$ be defined in the same way as $C(K)$ in Section **12** of CH(II) and let $(c^n_1, \cdots, c^n_{p_n})$ be a natural basis for $C_n(\tilde{K}) = H_n(\tilde{K}^n, \tilde{K}^{n-1})$. Let R, Γ, M mean the same as before, with $\gamma: \pi_1(K) \approx \Gamma$. Let $R(K)$ be the group ring of $\pi_1(K)$. Let $C_n(K) \subset M$ be a basic module of rank p_n $(n = 0, 1, \cdots)$, such that $C_i(K) \cap C_j(K) = 0$ if $i \neq j$. Let $(m^n_1, \cdots, m^n_{p_n})$ be the basis of $C_n(K)$ and let $k_n: C_n(\tilde{K}) \approx C_n(K)$ be defined by

$$k_n(r_1 c_{n1} + \cdots + r_{p_n} c^n_{p_n}) = (\gamma r_1) m^n_1 + \cdots + (\gamma r_{p_n}) m^n_{p_n}, \tag{10.1}$$

where $r_i \varepsilon R(K)$. Let

$$\partial_n = k_{n-1} \partial'_n k_n^{-1}: C_n(K) \to C_{n-1}(K), \tag{10.2}$$

where ∂' is the boundary operator in $C(\tilde{K})$. Obviously $\partial\partial = 0$ and $\partial r = r\partial$ $(r \varepsilon R)$. Therefore $C(K) = \{C_n(K)\}$, with $\partial = \{\partial_n\}$ as boundary operator, is a chain system and $k = \{k_n\}: C(\tilde{K}) \approx C(K)$ is an isomorphism associated with γ.

The arbitrariness in the definition of $C(K)$ consists of

a) the choice of the base point e^0,

(10.3) b) the choice of $\gamma: \pi_1(K) \approx \Gamma$,

c) the choices of the bases $\{c^n_i\}$ and of the basic modules $C_n(K)$.

Let another 0-cell $e_1^0 \varepsilon K^0$ be taken as base point and let $\tilde{K}_1$ and $C(\tilde{K}_1)$ be the corresponding universal covering complex and chain system. Let

$$\alpha: \pi_1(K, e_1^0) \approx \pi_1(K, e^0), \qquad \phi: \tilde{K}_1 \approx \tilde{K}$$

be the isomorphisms[27] determined by a path $(I, 0, 1) \to (K, e^0, e_1^0)$. Let $h: C(\tilde{K}_1) \approx C(\tilde{K})$ be the isomorphism induced by ϕ. Then h is obviously associated with α, and the same system, $C(K)$, is defined by

---

[26] Here we reserve the symbol $C(K)$ for a system in which $C_n(K)$ is a basic module in M, as defined by (10.1), (10.2) below. We no longer insist that the elementary chain $c^0 \varepsilon C_0(K)$, associated with a e-cell which covers e^0, shall be associated with the base point in K.

[27] We use the symbol $\approx$ to denote the relation of isomorphism between complexes as well as between groups and chain systems.

$$\gamma\alpha: \pi_1(K, e_1^0) \approx \Gamma, \qquad kh: C(\tilde{K}_1) \approx C(K),$$

as by α and k. The effect of choosing a different path, $(I, 0, 1) \to (K, e^0, e_1^0)$, is to replace α, h by $\theta_x\alpha$, xh, where $x \varepsilon \pi_1(K, e^0)$. Since

$$x(r_1c^n_1 + \cdot\cdot\cdot + r_{p_n}c^n_{p_n}) = (\theta_x r_1)xc^n_1 + \cdot\cdot\cdot + (\theta_x r_{p_n})xc^n_{p_n}$$

it follows that the resulting alterations, $\gamma\alpha \to \gamma\theta_x\alpha$ and $kh \to kxh$, are included in (b) and (c).

Let γ be replaced by $\gamma': \pi_1(K) \approx \Gamma$ and let $\theta = \gamma'\gamma^{-1}: \Gamma \approx \Gamma$. Then it follows from (10.1) and (10.2) that k_n is replaced by $s_\theta k_n$ and ∂ by $\partial^\theta = s_\theta \partial s_\theta^{-1}$. Therefore $C(K)$ is replaced by the conjugate system $C^\theta(K)$.

Any other natural basis for $C_n(\tilde{K})$ is of the form $(\pm x_1c^n_1, \cdot\cdot\cdot, \pm x_{p_n}c^n_{p_n})$, where $x_i \varepsilon \pi_1(K)$. Any other basic module of rank p_n is of the form PC_n, where $P: M \to M$ is a permutation. Therefore a change in (c) leads to a new system $C'(K) \approx C(K)$ (Σ).

Therefore $C(K)$ is determined up to a transformation, $C(K) \to C'(K)$, which is the resultant of a semi-linear transformation, $C(K) \to C^\theta(K)$, followed by a simple isomorphism $C^\theta(K) \approx C'(K)$ (Σ).

Let $K' \equiv K$ and let $k': C(\tilde{K}) \approx C(K')$ be defined in the same way as $k: C(\tilde{K}) \approx C(K)$, in terms of an isomorphism $\gamma': \pi_1(K') \approx \Gamma$. Let $g: C(\tilde{K}) \equiv C(\tilde{K}')$ and $\alpha: \pi_1(K) \approx \pi_1(K')$ be the chain equivalence and the isomorphism induced by a homotopy equivalence [28] $\phi: K \equiv K'$. Let

$$(10.4) \qquad f = k'gk^{-1}: C(K) \to C(K'), \qquad \theta = \gamma'\alpha\gamma^{-1}: \Gamma \approx \Gamma.$$

Then it may be verified that f is a chain equivalence associated with θ. We describe it as the chain equivalence *induced* by ϕ and we define $\tau(\phi) = \tau(f)$. Let $g^*: C(\tilde{K}) \equiv C(\tilde{K}')$ be the chain equivalence induced by a homotopic map $\phi^* \simeq \phi$ and let $f^* = k'g^*k^{-1}$. Then $g^* \simeq g$ and it follows that $f^* \simeq f$. Therefore $\tau(\phi^*) = \tau(\phi)$. Hence, and by the two preceding paragraphs, $\tau(\phi)$ depends only on the homotopy class, $\bar{\phi}$, of maps $K \to K'$, which contains ϕ, on γ, γ' and on the choice of base points [29] in K, K'. We define $\tau(\bar{\phi}) = \tau(\phi)$. Let γ and γ' be replaced by $\gamma_1: \pi_1(K) \approx \Gamma$ and $\gamma'_1: \pi_1(K') \approx \Gamma$. Let $\theta = \gamma_1\gamma^{-1}$, $\theta' = \gamma'_1\gamma'^{-1}$. Then k, k' are replaced by $s_\theta k$, $s_{\theta'}k'$ and f by $s_{\theta'}fs_\theta^{-1}$. Therefore it follows from (9.4) that $\tau(\bar{\phi})$ is replaced by $\theta'\tau(\bar{\phi})$.

Let us write $\tau \equiv \tau'$, where $\tau, \tau \subset \mathrm{T}$, if, and only if, $\tau' = \theta\tau$, where

[28] All our maps and homotopies of complexes will be cellular and it is always to be understood that a given map, $K \to K'$, carries e^0 into e'^0, where $e^0 \varepsilon K'^0$ are the base points.

[29] Actually $\tau(\phi)$ does not depend on the choice of base points since $\theta_\gamma\tau = \tau$ for any $\gamma \varepsilon \Gamma$, $\tau \varepsilon \mathrm{T}$.

$\theta : \Gamma \approx \Gamma$. Obviously $\tau \equiv \tau'$ is an equivalence relation and we shall describe the corresponding equivalence classes as *θ-classes.* It follows from the preceding paragraph that the θ-class, $\bar{\tau}(\bar{\phi})$, which contains $\tau(\bar{\phi})$, is uniquely determined by the homotopy class $\bar{\phi}$. We call it the *torsion* of $\bar{\phi}$, or of any map $\phi \, \varepsilon \, \bar{\phi}$.

We shall describe $\phi : K \equiv K'$ as a *simple* (*homotopy*) *equivalence,* and shall write $\phi : K \equiv K' \; (\Sigma)$, if, and only if, $\tau(\phi) = 0$. We shall say that K, K' are of the same *simple homotopy type,* and shall write $K \equiv K' \; (\Sigma)$, if, and only if, there is a simple homotopy equivalence $\phi : K \equiv K' \; (\Sigma)$.

Let $\phi' : K' \equiv K''$ and let $C(K'') \approx C(K'')$ be defined in the same way as $C(K)$ and $C(K')$, in terms of an isomorphism $\gamma'' : \pi_1(K'') \approx \Gamma$. Let ϕ' be associated with $\alpha' : \pi_1(K') \approx \pi_1(K'')$ and let $\theta' = \gamma''\alpha'\gamma'^{-1} : \Gamma \approx \Gamma$. Then it follows from (10.4) and (9.5) that

$$\tau(\phi'\phi) = \tau(\phi') + \theta'\tau(\phi). \tag{10.5}$$

Therefore, if ϕ, ϕ' are simple equivalences, so is $\phi'\phi$. Obviously $\tau(\psi) = 0$ if $\psi \simeq 1 : K \to K$. Therefore, taking $K'' = K$ and $\phi'\phi \simeq 1$, it follows from (10.5) that a homotopy inverse of a simple homotopy equivalence is itself a simple homotopy equivalence. Therefore $K \equiv K' \; (\Sigma)$ is an equivalence relation.

Let G_K be the aggregate of homotopy classes, $\bar{\phi}, \bar{\psi}, \cdots$, of homotopy equivalences, $\phi, \psi, \cdots$, of K into itself. Let $\bar{\psi}\bar{\phi} = \overline{\psi\phi}$. Then G_K, with this multiplication, is obviously a group. Let $e^0 \, \varepsilon \, K^0$ and $k : C(K) \approx C(K)$ be fixed and let $\mathfrak{G}_K$ be the sub-group of the groupoid $\mathfrak{G}$, which consists of the chain homotopy classes of chain equivalences $C(K) \equiv C(K)$. Let $f_\phi = f$, where f is given by (10.4), with $K' = K$, $\gamma' = \gamma$, $k' = k$, and let $\bar{f}_\phi \, \varepsilon \, \mathfrak{G}$ be the class which contains f_ϕ. Then $\bar{\phi} \to \bar{f}_\phi$ is obviously a homomorphism of G_K into $\mathfrak{G}_K$. Let $\tau_K : G_K \to \mathrm{T}$ be the map which is given by $\tau_K(\bar{\phi}) = \tau(\bar{\phi})$. It is the resultant of $\bar{\phi} \to \bar{f}_\phi$, followed by the crossed homomorphism $\bar{f} \to \tau(\bar{f})$. Therefore τ_K is a crossed homomorphism, in which G_K operates on T according to the rule $\bar{\phi}\tau = \bar{f}_\phi\tau$.

Let us take $\Gamma = \pi_1(K, p_0)$, where $p_0 \, \varepsilon \, K$, and let $\gamma : \pi_1(K, e^0) \approx \Gamma$ be the isomorphism determined by a path in K, which joins p_0 to e^0. Then the degree of arbitrariness in γ is that it may be replaced by $\theta\gamma$, where $\theta : \Gamma \approx \Gamma$ is an inner automorphism. In this case f_ϕ is replaced by $f_\phi{}^\theta$ and τ_K by $\theta\tau_K : G_K \to \mathrm{T}$. But $\theta\tau = \tau$, according to (2.11), whence $\theta\tau_K = \tau_K$. Therefore τ_K is uniquely determined by K when $\Gamma = \pi_1(K, p_0)$. For the reasons given in discussing (10.3), τ_K is independent of the choice of e^0.

Let K_0 be a connected sub-complex of K, which contains e^0 and is such

that $i:\pi_1(K_0) \approx \pi_1(K)$, where i is the injection homomorphism. Then $\tilde{K}_0 = p^{-1}K_0$ may be taken as the universal covering complex of K_0, where $p:\tilde{K} \to K$ is the covering map. Let $C_n(\tilde{K}_0) \subset C_n(\tilde{K})$ be the sub-module consisting of the n-chains carried by $\tilde{K}_0$ and let k_n mean the same as in (10.1). A natural basis for $C_n(\tilde{K}_0)$ is part of a natural basis for $C_n(\tilde{K})$ and it follows that $C(K_0)$, with

$$C_n(K_0) = k_n C_n(\tilde{K}_0), \tag{10.6}$$

is a sub-system of $C(K)$. Let $U = K - K_0$ and let us denote the residue system $C(K) - C(K_0)$ by $C(U) = C(K) - C(K_0)$. Let $\tilde{U} = p^{-1}U$ and let $C_n(\tilde{U}) \subset C_n(\tilde{K})$ be the sub-module consisting of the n-chains carried by $\tilde{U}$. Then obviously $C_n(U) = k_n C_n(\tilde{U})$.

When dealing with such a pair of complexes K and $K_0 \subset K$ we shall always assume that $C(K_0)$ is imbedded in $C(K)$ in the way described above.

Let $K_0 \subset K$, $L_0 \subset L$ be sub-complexes of given complexes K, L. Let $\phi: (K, K_0) \to (L, L_0)$ be a map such that $\phi \mid K - K_0$ is an isomorphism onto $L - L_0$ and let $\phi_0: K_0 \to L_0$ be the map which is induced by ϕ.

THEOREM [30] 11. *If ϕ_0 is a simple equivalence, so is ϕ.*

Let $h: C(K) \to C(L)$ and $f: C(K_0) \to C(L_0)$ be the chain mappings which are induced by ϕ and ϕ_0. Then it is obvious that $hC(K_0) \subset C(L_0)$ and that f is the chain mapping induced by h. Since $\phi \mid K - K_0$ is an isomorphism onto $L - L_0$ it is also obvious that $g: C(K - K_0) \approx C(L - L_0)$ (Σ), where g is the chain mapping induced by h. Therefore the Theorem follows from Theorem 10.

As an application of Theorem 11 let $\phi_0: K_0 \equiv L_0$, where L_0 consists of a single 0-cell, let L be formed from K by shrinking K_0 into the point L_0 and let $\phi: K \to L$ be the "identification map." Since $\pi_1(L_0) = 1$ it follows that ϕ_0 is a simple equivalence and so therefore is ϕ. In particular we can take $K_0 \subset K^1$ to be a tree containing K^0. Then L^0 consists of the single 0-cell L_0.

11. Combinatorial invariance.

In this section we prove:

THEOREM 12. *If K' is a sub-division of K the identical map, $i: K \to K'$, is a simple equivalence.*

---

[30] Cf. Theorem 12 in Section 8 of (I). Obviously $\phi_0: K_0 \equiv L_0(\Sigma)$ if $K_0 = e^0$, $L_0 = e'^0$, where $e^0 \in K^0$, $e'^0 \in L^0$ are the base points. In this case the theorem states that $\phi: K \equiv L(\Sigma)$ if $\phi: K \approx L$.

Let P be a given complex and $Q \subset P$ a sub-complex, which is a D. R. of P.

LEMMA 10. *If every circuit in $P - Q$ is contractible to a point in P, then the identical map, $Q \to P$, is a simple equivalence.*

We first prove the theorem, assuming the truth of the lemma. Let $\phi' : K' \approx L$, where L is a new complex, which does not meet K. Let $\phi = \phi' i : K \to L$. Then $i = \phi'^{-1}\phi$ and $\phi'^{-1} : L \equiv K' \ (\Sigma)$ according to Theorem 11. Therefore it is sufficient to prove that $\phi : K \equiv L \ (\Sigma)$. Let P be the mapping cylinder of ϕ. We regard P as $K \times I$, with $(x, 0) = x \ (x \varepsilon K)$ and $K \times 1$ sub-divided to form L. Let e_0 be a principal cell (i. e. one which is an open sub-set) of K and let $K_1 = K - e_0$. Proceeding by induction we define a sequence of sub-complexes

$$K = K_0, K_1, \cdots, K_n = K^{-1},$$

such that $K_{\lambda+1} = K_\lambda - e_\lambda$, where e_λ is a principal cell of K_λ. Let $P_\lambda = K \cup (K_\lambda \times I)$. Then $P_{\lambda+1}$ is a D. R.[31] of P_λ and $P_\lambda - P_{\lambda+1}$ is the point-set $e_\lambda \times (0, 1\rangle$, where $(0, 1\rangle$ is the half open interval $0 < t \leq 1$. Therefore $\pi_1(P_\lambda - P_{\lambda+1}) = 1$ and it follows from Lemma 10 that $i_\lambda : P_{\lambda+1} \equiv P_\lambda \ (\Sigma)$, where i_λ is the identical map. Therefore

$$j = i_0 \cdots i_{n-1} : P_n = K \equiv P \ (\Sigma).$$

Similarly $k : L \equiv P \ (\Sigma)$, where k is the identical map. Let $\psi : P \to L$ be given by $\psi(x, t) = \phi x$. Then $\psi : P \equiv L \ (\Sigma)$, since ψ is a homotopy inverse of k. Therefore $\phi = \psi j : K \equiv L \ (\Sigma)$ and Theorem 12 is proved.

It remains to prove Lemma 10. Since Q is a D. R. of P it is easily proved that the chain system $C(Q)$ is a D. R. of $C(P)$ and that a retraction $\psi : P \to Q$ induces a retraction $k : C(P) \to C(Q)$. We have to prove that k is a simple equivalence and this will follow from Theorem 4, Section **5**, when we have proved that $C(U) \equiv 0 \ (\Sigma)$, where $U = P - Q$.

Let $U_1, \cdots, U_m$ be the components of U, which are finite in number since U is the union of a finite number of (connected) cells. Let $\tilde{P}$ be the universal covering complex of P and let $\boldsymbol{p} : P \to \tilde{P}$ be the covering map. Let $\tilde{U}_\lambda$ be any component of $\boldsymbol{p}^{-1} U_\lambda$ and let $U^* = \tilde{U}_1 \cup \cdots \cup \tilde{U}_m$. Since P and $\tilde{P}$ are locally connected, U_λ and $\tilde{U}_\lambda$ are open sets. Let $e^n{}_1, \cdots, e^n{}_{a_n}$ be the n-cells in U. It follows from the condition on the circuits in U, which is satisfied a fortiori by the circuits in U_λ, that $\boldsymbol{p} \mid \tilde{U}_\lambda$ is a homeomorphism onto U_λ. Therefore $\boldsymbol{p} \mid U^*$ is a homeomorphism onto U. Therefore

[31] See Theorem 1. 4 (ii) in [16].

U^* contains precisely one, $\bar{e}^n_i$, of the cells in $\bar{P}$, which cover e^n_i. Let $c^n_i \varepsilon C_n(\bar{P})$ be the element which is represented by a characteristic map for $\bar{e}^n_i$. Then $(c^n_1, \cdots, c^n_{q_n})$ is a basis for $C_n(\bar{U})$, which is part of a natural basis for $C_n(\bar{P})$. Moreover $C_n(U^*)$, which consists of the n-chains carried by U^*, is the ordinary free Abelian group, which is freely generated by $c^n_1, \cdots, c^n_{q_n}$, without the help of the operators in $\pi_1(P)$.

Since each component of $\bar{U}$ is open it follows that no cell in $\bar{U} - U^*$ meets the closure of $\bar{e}^n_i$. Therefore

$$\partial c^n_i = \sum_{j=1}^{q_{n-1}} d_{ij}{}^n c_j{}^{n-1} + c'_i{}^{n-1}, \tag{11.1}$$

where $d_{ij}{}^n$ are integers and $c'_i{}^{n-1} \varepsilon C_n(\bar{Q})$ $(\bar{Q} = \mathbf{p}^{-1}Q)$. Let $m_i{}^n = k_n c_i{}^n$, where k_n means the same as in (10.1). Then it follows from (11.1) that $\partial\colon C(U) \to C(U)$ is given by $\partial m_i{}^n = \sum_{j=1}^{q_{n-1}} d_{ij}{}^n m_j{}^{n-1}$. Therefore $C(U) \equiv 0\ (\Sigma)$, by the corollary to Lemma 4, in Section **6**. This proves Lemma 10.

12. Lens spaces. By way of an example let A, B be the chain systems determined by lens spaces of types (m, p), (m, q), where m is the order of their fundamental groups and $q \equiv k^2 p(m)$. That is to say, A, B play the part of $C(K)$ in Section **10** and $\pi_1(P)$, $\pi_1(Q)$ in Section **15** of CH(II) are both replaced by Γ. The generators ξ, η in CH(II) are replaced by a generator $\gamma \varepsilon \Gamma$ and we denote the integer r by h, to avoid confusion with $r \varepsilon R$. Otherwise the notations will be the same as in CH(II). Thus $\partial\colon A \to A$ and $\partial\colon B \to B$ are given by

$$\begin{cases} \partial a_1 = (\gamma - 1)a_0, & \partial a_2 = \sigma_m(\gamma)a, \quad \partial a_3 = (\gamma^p - 1)a_2 \\ \partial b_1 = (\gamma - 1)b_0, & \partial b_2 = \sigma_m(\gamma)b_1, \quad \partial b_3 = (\gamma^q - 1)b_2, \end{cases} \tag{12.1}$$

where $\sigma_t(\gamma^s) = 1 + \gamma^s + \cdots + \gamma^{(t-1)s}$.

The Reidemeister-Franz torsion in A and B is r_A and r_B, where

$$r_A = (\gamma - 1)(\gamma^p - 1), \qquad r_B = (\gamma - 1)(\gamma^q - 1). \tag{12.2}$$

Let $\theta\colon \Gamma \approx \Gamma$ be given by $\theta\gamma = \gamma^k$ and let $u\colon A \to B$, $v\colon B \to A$ be the chain mappings, associated with θ and with θ^{-1}, which are given by $ua_n = b_n$, $vb_n = a_n$ if $n = 0$ or 3 and

$$\begin{cases} ua_1 = \sigma_k(\gamma)b_1, & ua_2 = \sigma_k(\gamma^{kp})b_2 \\ vb_1 = \sigma_l(\gamma)a_1, & vb_2 = \sigma_l(\gamma^{lq})a_2, \end{cases} \tag{12.3}$$

where $kl = 1 + mh$. As shown in CH(II), $vu - 1 = \partial\eta + \eta\partial$, $uv - 1$

$= \partial\eta + \eta\partial$, where $\eta a_n = \eta b_n = 0$ if $n \neq 1$ and $\eta a_1 = ha_2$, $\eta b_1 = hb_2$. Notice that $\eta\eta = 0$ and $\eta s_\theta = s_\theta \eta$, since a_n, b_n are generators, m_{i_n}, m_{j_n}, of M.

It follows from (12.1) that

$$\text{(12.4)} \qquad \partial^\theta a_1 = (\gamma^k - 1)a_0, \quad \partial^\theta a_2 = \sigma_m(\gamma)a_1, \quad \partial^\theta a_3 = (\gamma^{kp} - 1)a_2,$$

and from (12.3) that

$$\text{(12.5)} \qquad s_\theta v b_1 = \sigma_l(\gamma^k)a_1, \qquad s_\theta v b_2 = \sigma_l(\gamma^q)a_2,$$

since $kl \equiv 1(m)$. We proceed to calculate $\tau(u) = \tau\{C(us_\theta^{-1})\}$, where $C(f)$ means the same as in Section **8**. Let $C = C(us_\theta^{-1})$ and let $a'_n = \alpha a_n$, where α means the same as in (8.2). Then (b_n, a'_{n-1}) is a basis for C_n $(a'_{-1} = b_4 = 0)$. Since $us_\theta^{-1}a_n = ua_n$ it follows from (8.2), (12.3) and (12.4) that $\partial : C \to C$ is given by

$$\begin{aligned}
\partial b_1 &= (\gamma - 1)b_0, & \partial a'_0 &= b_0 \\
\partial b_2 &= \sigma_m(\gamma)b_1, & \partial a'_1 &= \sigma_k(\gamma)b_1 - (\gamma^k - 1)a'_0 \\
\partial b_3 &= (\gamma^q - 1)b_2, & \partial a'_2 &= \sigma_k(\gamma^{kp})b_2 - \sigma_m(\gamma)a'_1 \\
& & \partial a'_3 &= b_3 - (\gamma^{kp} - 1)a'_2.
\end{aligned}$$

It is easily verified that $\mu = 0$, where μ is given by (8.6), and similarly that $s_\theta v\eta = \eta s_\theta v$. Since $\eta\eta = 0$ a straightforward caluculation shows that $\delta\delta = 0$, where δ is given by (8.7), with $\xi = \delta$. It follows from (8.7) and (12.5) that $\delta b_0 = a'_0$ and

$$\begin{aligned}
\delta b_1 &= -hb_2 + \sigma_l(\gamma^k)a'_1, & \delta a'_0 &= 0 \\
\delta b_2 &= \sigma_l(\gamma^q)a'_2, & \delta a'_1 &= ha'_2 \\
\delta b_3 &= a'_3, & \delta a'_2 &= 0.
\end{aligned}$$

Let $D_0 = C_0 + C_2 + C_4$, $D_1 = C_1 + C_3$, as in (6.8) with $m = \dim C - 1 = 3$. Then D_0, D_1 have (b_0, a'_1, b_2, a'_3), (a'_0, b_1, a'_2, b_3) as bases and $\Delta_0 = \partial + \delta : D_0 \to D_1$ is given by

$$\begin{aligned}
\Delta_0 b_0 &= a'_0 \\
\Delta_0 a'_1 &= -(\gamma^k - 1)a'_0 + \sigma_k(\gamma)b_1 + ha'_2 \\
\Delta_0 b_2 &= \sigma_m(\gamma)b_1 + \sigma_l(\gamma^q)a'_2 \\
\Delta_0 a'_3 &= -(\gamma^{kp} - 1)a'_2 + b_3.
\end{aligned}$$

Let $f_0 : D_0 \approx D_0$, $f_1 : D_1 \approx D_1$ be the simple automorphisms which are given by

$$\begin{aligned}
f_0 a'_1 &= a'_1 + (\gamma^k - 1)b_0, & f_0 c &= c & (c &= b_0, b_2, a'_3) \\
f_1 b_3 &= b_3 + (\gamma^{kp} - 1)a'_2, & f_1 c &= c & (c &= a'_0, b_1, a'_2).
\end{aligned}$$

Then $\Delta'_0 = f_1\Delta_0 f_0 : D_0 \to D_1$ is given by $\Delta'_0 b_0 = a'_0$, $\Delta'_0 a'_3 = b_3$ and

$$\Delta'_0 a'_1 = \sigma_k(\gamma) b_1 + h a'_2$$
$$\Delta'_0 b_2 = \sigma_m(\gamma) b_1 + \sigma_l(\gamma^q) a'_2.$$

Let i, j be any integers and let $\rho = (i, j)$. If $i \leq j$ we have, writing $\sigma_s = \sigma_s(\gamma)$ $(\sigma_0 = 0)$,

$$\sigma_j - \gamma^{j-i}\sigma_i = \sigma_{j-i}.$$

Therefore it follows by induction on $i + j$ that the matrix $[\sigma_i, \sigma_j]$ can be reduced to $[\sigma_\rho, 0]$ by a sequence of transformations of the form

$$[\sigma_i, \sigma_j] \to [\sigma_i, \sigma_j - \gamma^s\sigma_i] \text{ or } [\sigma_i - \gamma^s\sigma_j, \sigma_j]$$

followed, if necessary, by $[0, \sigma_\rho] \to [\sigma_\rho, 0]$. Since $(k, m) = 1$ it follows that

$$\begin{bmatrix} \sigma_k(\gamma) & h \\ \sigma_m(\gamma) & \sigma_l(\gamma^q) \end{bmatrix} \to \begin{bmatrix} 1 & r \\ 0 & t \end{bmatrix} \qquad (r, t \subset R)$$

by such elementary transformations of the rows. Since these transformations alter the determinant by a factor ± 1, at most, we have $t = \pm d$, where

$$d = \sigma_k(\gamma)\sigma_l(\gamma^q) - h\sigma_m(\gamma).$$

I say that

$$(12.6) \qquad d r_B = \theta r_A,$$

where r_A, r_B are given by (12.2). For let χ be the homomorphism of R into the complex field, which is given by $\chi r = \omega$, where $\omega^m = 1$. Then $\chi(dr_B) = \chi(\theta r_A) = 0$ if $\omega = 1$, and if $\omega \neq 1$ we have

$$\chi(dr_B) = [(\omega^k - 1)(\omega^{lq} - 1)]/[(\omega - 1)(\omega^q - 1)](\omega - 1)(\omega^q - 1)$$
$$= (\omega^k - 1)(\omega^{lq} - 1) = \chi(\theta r_A).$$

Therefore $\chi(dr_B - r_A) = 0$ and (12.6) follows from the orthogonality relations between the group characters. Therefore $\tau(u)$ is essentially the same as the inverse of the element π in Lemma 5 on p. 1209 of [3].

13. Formal deformations. As in CH(II) let I^n be the n-cube in Hilbert space, which is given by $0 \leq t_1, \cdots, t_n \leq 1$, $t_i = 0$ if $i > n$, with $I^0 = (0, 0, \cdots)$. Let

$$E_0^{n-1} = \partial I^n - (I^{n-1} - \partial I^{n-1}) \qquad (n \geq 1).$$

Let e^n be a principal cell of a complex K and let e^{n-1} be a principal

cell of $K - e^n$ $(n \geq 1)$. We shall describe the transformation $K \to K_1 = K - e^n - e^{n-1}$ as an *elementary contraction* if, and only if, e^n has a characteristic map, $f: I^n \to \bar{e}^n$, such that $fE_0^{n-1} \subset K_1$ and $f \mid I^{n-1}$ is a characteristic map for e^{n-1}. The inverse, $K_1 \to K$, of an elementary contraction, $K \to K_1$, will be called an *elementary expansion.* An elementary expansion may also be defined as follows. Let E^n be an n-element, which is disjoint from a given complex, K, and let E^{n-1} be a hemisphere [32] of ∂E^n. Let $e^n = E^n - \partial E^n$, $e^{n-1} = \partial E^n - E^{n-1}$ and let $f: (E^{n-1}, \partial E^{n-1}) \to (K^{n-1}, K^{n-2})$ be an arbitrary map. Let $K_1 = K \cup e^{n-1} \cup e^n$ be the complex formed by identifying each point $p \varepsilon E^{n-1}$ with $fp \varepsilon K^{n-1}$. Then $K \to K_1$ is obviously an elementary expansion.

Either an elementary expansion or an elementary contraction will be called an *elementary deformation.* The resultant, $K_0 \to K_r$, of a finite sequence of elementary deformations,

$$K_i \to K_{i+1} \qquad (i = 0, \cdots, r-1), \tag{13.1}$$

will be called a *formal deformation.* We also include the identical transformation, $K \to K$, of any complex, among the formal deformations. We shall denote a formal deformation by $D: K_0 \to K_r$, and $K_r = DK_0$ will mean that K_r is obtained from K_0 by a formal deformation D. If each $K_i \to K_{i+1}$ is an elementary expansion (contraction) then $K_0 \to K_r$ will be called an *expansion* (*contraction*) and we shall say that K_0 *expands* (*contracts*) into K_r. If $D: K_0 \to K_r$ is the resultant of the sequence (13. 1), then the resultant of the sequence $K_{i+1} \to K_i$ is the formal deformation, $D^{-1}: K_r \to K_0$, *inverse* to D.

Let $D: K_0 \to K_1$ be an elementary contraction. Obviously E_0^{n-1} is a D. R. of I^n. Therefore [33] K_1 is a D. R. of K_0 and $[D]$ will denote the homotopy class of maps, $K_0 \to K_1$, which contains a retraction. If $D: K_0 \to K_1$ is an elementary expansion then $[D]$ will denote the homotopy class of maps, $K_0 \to K_1$, which contains the identity. Let $D = D_{r-1} \cdots D_0 : K_0 \to K_r$, where D_i stands for (13. 1). We define $[D]$ by $[D] = [D_{r-1}] \cdots [D_0]$. If $1: K \to K$ is the identical formal deformation then $[1]$ will denote the homotopy class of the identical map $K \to K$. It is obvious that, if $D: K \to K'$ and $D': K' \to K''$ are formal deformations, then $K'' = D'DK$ and

$$[D'D] = [D'][D]. \tag{13.2}$$

Also it is easily verified that $[D^{-1}][D] = [1]$.

---

[32] I. e. the image of E_0^{n-1} in some homeomorphism $\partial I^n \to \partial E^n$.

[33] See Lemma 2 in Section 4 of [4].

Let L be a sub-complex of K, which may be empty. By a formal deformation, $D: K \rightarrow K'$, rel. L, we shall mean the resultant of a sequence of elementary deformations, none of which removes a cell of L. By $K' = DK$, rel. L, we shall mean that K' is the image of K in such a formal deformation, D. If $K' = DK$, rel. L, then $[D]$ obviously contains at least one map $\phi_0: K \rightarrow K'$, rel. L, where rel. L means that $\phi_0 y = y$ if $y \varepsilon L$. We restrict $[D]$ to maps, $\phi: K \rightarrow K'$, rel. L, such that $\phi \simeq \phi_0$, rel. L.

We shall describe a map $\phi: K \rightarrow K'$, rel. L, as a *restricted equivalence*, rel. L, if, and only if, $K' = DK$, rel. L, and $\phi \varepsilon [D]$. We shall write $K \equiv K' (\Sigma)$, rel. L, if, and only if, there is a simple equivalence $\phi: K \equiv K' (\Sigma)$, which is rel. L.

THEOREM 13. *$K' = DK$, rel. L, if, and only if, $K \equiv K' (\Sigma)$, rel. L, in which case the restricted equivalences, $K \rightarrow K'$, rel. L, are the same as the simple equivalences, $K \rightarrow K'$, rel. L.*

In order to prove this we shall need some lemmas, which are proved in the following section.

14. Lemmas on formal deformations. Let K, K' be complexes, with a common sub-complex, L, which may be empty, and let $(K - L) \cap K' = 0$. Let $\phi: K \rightarrow K'$ be a map rel. L. By the *mapping cylinder* of ϕ we mean the complex, P, which is formed from[34] $K \times I$ by identifying $(x, 0)$ with x, $(x, 1)$ with ϕx and $y \times I$ with y, for each point $x \varepsilon K$ and $y \varepsilon L$. Let

$$\phi_r: \pi_r(K) \approx \pi_r(L) \qquad (r = 1, \cdots, m), \tag{14.1}$$

where ϕ_r is the homomorphism induced by ϕ. Then the argument[35] used in the case $L = 0$ shows that

$$i_r: \pi_r(K) \approx \pi_r(P) \qquad (r = 1, \cdots, m), \tag{14.2}$$

where i_r is the injection. Therefore it follows from the exactness of the sequence

$$\pi_r(K) \rightarrow \pi_r(P) \rightarrow \pi_r(P, K) \rightarrow \pi_{r-1}(K) \rightarrow \pi_{r-1}(P),$$

that $\pi_r(P, K) = 0$ for $r = 1, \cdots, m$, where $\pi_1(P, K) = 0$ means that i_1 is onto $\pi_1(P)$.

[34] After replacing $K \times I$ by a homeomorph, if necessary, we assume that it has no point in common with K or K'.

[35] See Section 3 of [5].

LEMMA 11. *P contracts into K'.*

Let $K = K_0 \cup e^n$, where e^n is a principal cell in $K - L$. Then $P = P_0 \cup e^n \cup e^{n+1}$, where $e^{n+1} = e^n \times (0,1)$ and P_0 is the mapping cylinder of $\phi \mid K_0 : K_0 \to K'$. Let $f : I^n \to \bar{e}^n$ be a characteristic map for e^n. Then $g : I^{n+1} \to \bar{e}^{n+1}$, given by $g(t_1, \cdots, t_n, t) = \{f(t_1, \cdots, t_n), t\}$, is obviously a characteristic map for e^{n+1} and $gE_0{}^n \subset P_0$ and $gx = fx$ if $x \varepsilon I^n$. Therefore P contracts into P_0. Therefore the lemma follows by induction on the number of cells in $K - L$.

Let $\psi : P \to K'$ be given by

$$\psi \mid K' = 1, \qquad \psi(x, t) = \phi x \qquad (x \varepsilon K).$$

Since $\psi \mid K' = 1$ and since any two retractions $P \to K'$ are homotopic to each other, we have $\psi \varepsilon [D]$, where $D : P \to K'$ is any contraction. Also $\phi = \psi i$, where $i : K \to P$ is the identical map. If $K = D_1 P$, rel. K, then $i \varepsilon [D_1^{-1}]$. Hence, and from (13.2), we have the corollary:

COROLLARY. *If $K = D_1 P$, rel. K, then $\phi \varepsilon [DD_1^{-1}]$.*

Let K, K' and L be as in Lemma 11, except that $K - L$ and $K' - L$ may now have points in common. Let $\phi : (K, L) \approx (K', L)$, rel. L.

LEMMA 12. *ϕ is a restricted equivalence, rel. L.*

First let $K \cap K' = L$ and let P be the mapping cylinder of ϕ. Then P may also be regarded as the mapping cylinder of ϕ^{-1} and the Lemma follows from Lemma 11 and its corollary.

If $K' - L$ meets K we replace the points in $K' - L$ by new ones, thus forming a complex K'', such that $\phi' : K'' \approx K'$, rel. L, and $K \cap K'' = K' \cap K'' = L$. By what we have already proved, ϕ' and $\phi'^{-1}\phi : K \approx K''$ are restricted equivalences, rel. L. Therefore it follows from (13.2) that ϕ is a restricted equivalence, rel. L, and the lemma is proved.

Let K_0, K_1 be complexes with a common sub-complex, K, and let $K_i = K \cup e_i{}^n$ $(i = 0, 1; n \geq 1)$. Let $f_i : I^n \to e_i{}^n$ be a characteristic map for $e_i{}^n$ in K_i.

LEMMA 13. *If $f_0 \mid \partial I^n \simeq f_1 \mid \partial I^n$ in K, then $K_1 = DK_0$, rel. K.*

First assume that $e_0{}^n \cap e_1{}^n = 0$ and unite K_0, K_1 in the complex $K^* = K_0 \cup K_1$. Let $g_t : \partial I^n \to K$ be a homotopy of $g_0 = f_0 \mid \partial I^n$ into $g_1 = f_1 \mid \partial I^n$ and let $f : \partial I^{n+1} \to K^*$ be given by

$$f(t_1,\cdots,t_n,i)=f_i(t_1,\cdots,t_n) \quad . \qquad \{i=0,1;(t_1,\cdots,t_n)\,\varepsilon\, I^n\}$$
$$f(s_1,\cdots,s_n,t)=g_t(s_1,\cdots,s_n) \qquad \{t\,\varepsilon\, I;(s_1,\cdots,s_n)\,\varepsilon\,\partial I^n\}.$$

We attach a new cell, e^{n+1}, to K^* by means of the map[36] f, thus forming a complex, $L=K^*\cup e^{n+1}$, in which e^{n+1} has a characteristic map, $h:I^{n+1}\to\bar{e}^{n+1}$ such that $h\mid\partial I^{n+1}=f$. Since $h(x,0)=f_0x$ $(x\,\varepsilon\, I^n)$ and[37] $hE_0{}^n\subset L-e_0{}^n$ it follows that $L\to K_1$ is an elementary contraction. Similarly $L\to K_0$ is an elementary contraction. Therefore $K_0\to L\to K_1$ is a formal deformation, rel. K.

If $e_0{}^n\cap e_1{}^n\neq 0$ we attach a new cell, $e'_0{}^n$, to K, by means of the map $f_0\mid\partial I^n$, taking care that $e'_i{}^n\cap e_i{}^n=0$ $(i=0,1)$. Then $K_0\to K\cup e'_0\to K_1$ is a formal deformation, rel. K, and the lemma is proved.

Let P be a given complex, let $P_0\subset P$ be a sub-complex and let $k_n(P-P_0)$ denote the number of n-cells in $P-P_0$. Let K be a sub-complex of P_0 and let $D_0:P_0\to Q_0$ be a formal deformation, rel. K. We shall describe a formal deformation, $D:P\to Q$, rel. K, as an *extension* of D_0 if, and only if, Q_0 is a sub-complex of Q and

$$k_n(Q-Q_0)=k_n(P-P_0) \qquad (n=0,1,\cdots).$$

LEMMA 14. *$D_0:P_0\to Q_0$ has an extension $D:P\to Q$.*

Let $D:P\to Q$ be an extension of D_0 and let $D':Q\to Q'$, rel. K, be an extension of a formal deformation $D'_0:Q_0\to Q'_0$. Then $D'D:P\to Q'$ is obviously an extension of $D'_0D_0:P_0\to Q'_0$. Let $P_1\subset P$ be a sub-complex, which contains P_0. Let $D_1:P_1\to Q_1$, rel. K, be an extension of D_0 and let $D:P\to Q$ be an extension of D_1. Then D is obviously an extension of D_0. Therefore the Lemma will follow by a double induction on the number of elementary deformations in D_0 and on the number of cells in $P-P_0$ when we have proved it in case D_0 is an elementary deformation and $P-P_0$ is a single cell.

Let $P=P_0\cup e^n$ and let D_0 be an elementary expansion $D_0:P_0\to Q_0=P_0\cup e^{p-1}\cup e^p$. If e^n has a point in common with $e^{p-1}\cup e^p$ we apply a preliminary formal deformation, $P\to P'$, rel. P_0, as in Lemma 13, so as to replace e^n by a cell which is disjoint from $e^{p-1}\cup e^p$. Then P' and Q_0 may be united in a complex

$$Q=P'\cup e^{p-1}\cup e^p$$

[36] That is to say, we attach an $(n+1)$-element, E^{n+1} $(e^{n+1}=E^{n+1}-\partial E^{n+1})$, to K^* by means of the map fh': $\partial E^{n+1}\to K^*$, where h': $\partial E^{n+1}\to\partial I^{n+1}$ is a homomorphism.

[37] We recall that $E_0{}^{n+1}=\partial I^{n+1}-(I^n-\partial I^n)$.

and $P \to P' \to Q$ is an extension of D_0.

Let D_0 be an elementary contraction,

$$D_0 : P_0 \to Q_0 = P_0 - e^p - e^{p-1},$$

and let $f : I^n \to \bar{e}^n$ be a characteristic map for e^n. Since Q_0 is a D. R. of P_0 there is a map, $f' : \partial I^n \to Q_0$, which is homotopic, in P_0, to $f \mid \partial I^n$. We attach a new cell, e'^n, to P_0 by means of the map f', thus forming a complex $P' = P_0 \cup e'^n$. Then $P' = D'P$, rel. P_0, by Lemma 13. Since $\partial e'^n \subset Q_0$ it follows that $P' \to Q = P' - e^p - e^{p-1}$ is an elementary contraction and $P \to P' \to Q$ is an extension of D_0. This proves the lemma.

Let K be a (connected) sub-complex of P such that $\pi_n(P, K) = 0$ $(n = 1, \cdots, r)$.

LEMMA 15. *There is a formal deformation $D : P \to Q$, rel. K, such that $k_n(Q - K) = 0$ if $n \leq r$ and $k_n(Q - K) = k_n(P - K)$ if $n > r + 2$.*

Let $0 \leq p \leq r$ and assume that, if $p > 0$, then $k_n(P - K) = 0$ for $n = 0, \cdots, p - 1$. For the sake of clarity we consider the case $p = 0$ separately. Let $p = 0$ and let $e_1^0, \cdots, e_k^0$ be the 0-cells in $P - K$. Since P is connected there is a map

$$g_i : (E_0^1, I^0, E_0^0) \to (P^1, e_i^0, K^0).$$

Let $E_1^2, \cdots, E_k^2$ be a set of 2-elements, which are disjoint from P and from each other. Let $h_i : \partial E_i^2 \to \partial I^2$ be a homeomorphism and let

$$e_i^2 = E_i^2 - \partial E_i^2, \qquad e_i^1 = h_i^{-1}(I^1 - \partial I^1) = h_i^{-1}(\partial I^2 - E_0^1).$$

We attach E_i^2 to P by means of the map $g_i h_i : h_i^{-1} E_0^1 \to P^1$, thus forming a complex $P^* = P \cup \bigcup_i (e_i^1 \cup e_i^2)$. Then $P \to P^*$ is an expansion. The complex $K^* = K \cup \bigcup_i (e_i^0 \cup e_i^1)$ contracts into K. By Lemma 14 there is an extension, $P^* \to M_0$, rel. K, of the contraction $K^* \to K$. Then

$$k_0(M_0 - K) = k_0(P^* - K^*) = 0$$
$$k_n(M_0 - K) = k_n(P^* - K^*) = k_n(P - K) \qquad (n > 2).$$

Thus we have eliminated the 0-cells from $P - K$ at the expense of introducing $k_0(P - K)$ new 2-cells.

Now let $p > 0$, let $e_1^p, \cdots, e_k^p$ be the p-cells in $P - K$ and let $f_i : I^p \to \bar{e}_i^p$ be a characteristic map for e_i^p. Since $k_{p-1}(P - K) = 0$, whence

$P^{p-1} = K^{p-1}$, we have $f_t \partial I^p \subset K$. Since $\pi_p(P, K) = 0$ and since I^p, $E_0{}^p$ are two hemispheres of $\partial E_0{}^{p+1} = \partial I^{p+1}$, the map f_t can be extended to a map,

$$(14.3) \qquad g_t : (E_0{}^{p+1}, E_0{}^p) \to (P^{p+1}, K^p).$$

It now follows, in exactly the same way as when $p = 0$, that there is a complex $M_p = D_p P$, rel. K, such that $k_n(M_p - K) = 0$ if $n \leq p$ and $k_n(M_p - K) = k_n(P - K)$ if $n > p + 2$. Therefore the lemma follows by induction on p.

15. Proof of Theorem 13. Let $D : K \to K_1 = K \cup e^{n-1} \cup e^n$, $(n \geq 1)$ be an elementary expansion. Then $i \,\varepsilon\, [D]$, where $i : K \to K_1$ is the identity. Also K is a D. R. of K_1 and $K_1 - K$ is simply connected. Therefore it follows from Lemma 10 that $i : K \to K_1$ is a simple equivalence. Since $k : K_1 \to K$ is a homotopy inverse of i if $k \,\varepsilon\, [D^{-1}]$, it follows that k is also a simple equivalence. Therefore $\phi : K \to DK$ is a simple equivalence, rel. L, if $\phi \,\varepsilon\, [D]$, where D is any elementary deformation, rel L. Therefore it follows from an inductive argument that $\phi : K \equiv K'$ (Σ), rel. L, if $K' = DK$, rel. L, and $\phi \,\varepsilon\, [D]$.

Conversely, let $\phi : K \equiv K'$ (Σ), rel. L. Then it follows from Theorem 11 and Lemma 12 that we may, without loss of generality, replace K' by K'', where $K'' \approx K'$, rel. L. Therefore we assume that $K \cap K' = L$ and also that $e^0 = e'^0 \,\varepsilon\, L$, where $e^0 \,\varepsilon\, K^0$, $e'^0 \,\varepsilon\, K'^0$ are the base points, thus excluding the case $L = 0$. Let P, with base point e^0, be the mapping cylinder of ϕ. Since $\phi : K \equiv K'$ the relations (14.1), (14.2) hold for every $n \geq 1$. Also K' is a D. R. of P. Therefore $i'_1 : \pi_1(K') \approx \pi_1(P)$, where i'_1 is the injection. Therefore $C(K)$, $C(K')$ are sub-systems of $C(P)$, according to the convention (10.6). Let $j : C(L) \to C(P)$ be the chain mapping induced by the identical map $L \to P$. Then $A = jC(L) = C(K) \cap C(K')$. Let $h : C(K) \to C(K')$ be the chain mapping induced by ϕ. Then h is obviously rel. A, since ϕ is rel. L. It may be verified [38] that $C(P)$ is the mapping cylinder of h, as defined in the paragraph containing (8.4). Since, by hypothesis, h is a simple equivalence we have

$$(15.1) \qquad C(P - K) \equiv 0 \ (\Sigma).$$

according to (8.4).

[38] See Section **14** of CH(II), with $\gamma c_\lambda{}^n = 0$ if $c_\lambda{}^n$ corresponds to a cell in $L \subset K$ (not the L in CH(II)).

It follows from (14.1) that $\pi_r(P, K) = 0$ for every $r \geqq 1$. Let

(15.2) $q = \mathrm{Max}\{\dim(P - K), 3\} = \mathrm{Max}\{\dim(K - L) + 1, \dim(K' - L), 3\}$.

Then it follows from Lemma 15 that there is a complex $Q = D_1 P$, rel. K, such that

$$(15.3) \qquad Q = K \cup e_1^{q-1} \cup \cdots \cup e_s^{q-1} \cup e_1^q \cup \cdots \cup e_t^q.$$

By the first part of the Theorem, with K, K', L replaced by P, Q, K, we have $C(Q) \equiv C(P)$ (Σ), rel. $C(K)$. Therefore it follows from Theorem 4(a) and (15.1) that $C'' = C(Q - K) \equiv C(P - K) \equiv 0$ (Σ), whence $\partial''_q : C''_q \approx C''_{q-1}$ (Σ), by the corollary to Theorem 5. Therefore $s = t$. Let

$$\partial''_q m_i^q = \sum_{j=1}^{s} d_{ij} m_j^{q-1} \qquad (d_{ij} \varepsilon R),$$

where $(m_1^r, \cdots, m_s^r)$ is the basis of C''_r $(r = q - 1, q)$. Since $\tau(\partial''_q) = 0$ the matrix $\boldsymbol{d} = [d_{ij}]$ can be annihilated by an expansion

$$\boldsymbol{d} \rightarrow \begin{bmatrix} \boldsymbol{d} & 0 \\ 0 & \mathbf{1}_k \end{bmatrix},$$

followed by a sequence of elementary transformations of the form (2.12), followed by a contraction $\mathbf{1}_{s+k} \rightarrow \mathbf{1}_0$, where $\mathbf{1}_0$ is the empty matrix. Since $q - 1 \geqq 2$ it follows from arguments on pp. 289, 290 of [1], with minor alterations,[39] that the transformation $\boldsymbol{d} \rightarrow \mathbf{1}_0$ can be "copied geometrically" by a formal deformation $Q \rightarrow K$, rel. K. Therefore $K = D_2 P$, rel. K, and the Theorem follows from the corollary to Lemma 11.

Let us describe r as the order of an elementary expansion $K \rightarrow K_1 = K \cup e^{r-1} \cup e^r$, and also of its inverse, $K_1 \rightarrow K$. Let $\phi : K \equiv K'$ (Σ), rel. L and let $K^p, K'^p \subset L$, for some $p \geqq -1$. Then the following addendum to Theorem 13 is implicit in the proofs of Lemmas 11-15 and of Theorem 13.

Addendum. *$K' = DK$, rel. L, and $\phi \,\varepsilon\, [D]$, where D is the resultant of elementary deformations, whose orders lie between $p + 2$ and $q + 1$ inclusive, where q is given by* (15.2).

This addendum has the following application. It follows from Theorems 11 and 13 that, by means of a formal deformation, we can reduce a given complex to one which has a given point, e^0, as its only 0-cell. It is some-

---

[39] Let $r(a_1, \cdots, a_k)$ mean the same as in Theorem 19 of [1], with $K^* = Q^{q-1}$ and $k = s$. Then $m_j^{q-1} \rightarrow a_j$ determines an isomorphism $C''_{q-1} \approx r(a_1, \cdots, a_k)$.

times convenient to restrict ourselves to a class of complexes, all of which have the same point, e^0, as their only 0-cell. Let K_0, K_r be two such complexes and let $K_r = DK_0$. Then it follows from Theorem 13 and its addendum, with $p = 0$, that $[D] = [D']$, where $D' : K_0 \rightarrow K_r$ is the resultant of elementary deformation, $K_i \rightarrow K_{i+1}$, whose orders exceed 1. Therefore $K_j{}^0 = e^0$ for each $j = 0, \cdots, r$.

16. *n*-types. By a *cluster of n-spheres, attached to a space X,* at a point $x_0 \in X$, we shall mean a set of n-spheres, $\{S_i{}^n\}$, such that $X \cap S_i{}^n = x_0$ and $S_i{}^n - x_0$ does not meet $S_j{}^n$ if $i \neq j$. If X is a complex and $x_0 \in X^{n-1}$, then $X \cup \{S_i{}^n\}$ is the complex $X \cup \{e_i{}^n\}$, where $e_i{}^n = S_i{}^n - x_0$. At this stage we assume that, X being a finite complex, the number of n-spheres in a cluster attached to X is finite.

Let K^n, L^n be complexes of at most n dimensions $(n > 1)$ and let $\phi : K^n \rightarrow L^n$ be an $(n-1)$-homotopy equivalence, as defined in Section 2 of CH(I).

THEOREM 14. *There is a simple equivalence,*

$$\psi : K^n \cup \{S_{1i}{}^n\} \equiv L^n \cup \{S_{2j}{}^n\} \quad (\Sigma),$$

such that $\psi x = \phi x$ *if* $x \in K^{n-1}$, *where* $\{\ \ _i{}^n\}$, $\{S_{2j}{}^n\}$ *are clusters of n-spheres attached to* K^{n-1}, L^{n-1}.

Assuming that $K^n \cap L^n = 0$, let P be the mapping cylinder of ϕ. Then P^n is the union of K^n, L^n and the cells $e^r \times (0, 1)$, where $e^r \in K^{n-1}$. We attach a cluster of n-spheres,

$$\{S_{2\rho}{}^n\} = S_{21}{}^n \cup \cdots \cup S_{2k}{}^n,$$

to a 0-cell $e'^0 \in L^0$, where k is to be determined later. Using Lemma 13, we transfer these over P^n to a 0-cell of K^n, so that they become a cluster, $\{S_\rho{}^n\}$, attached to K^{n-1}. Since $\phi : K^n \equiv_{n-1} L^n$ it follows that (14.1) and (14.2) are satisfied with $m = n - 1$. Therefore

(16.1)
a) $\pi_r(P^n, K^n) = 0$ *if* $r = 1, \cdots, n-1$,
b) *the injection,* $\pi_{n-1}(K^n) \rightarrow \pi_{n-1}(P^n)$, *is an isomorphism (onto).*

These conditions are obviously satisfied by K^* and P^*, where

$$K^* = K^n \cup \{S_\rho\}, \qquad P^* = P^n \cup \{S_\rho{}^n\}.$$

Therefore it follows from (16. 1a) and Lemma 15 that there is a formal deformation

$$D_1 : P^* \to Q = K^* \cup e_1^{n-1} \cup \cdots \cup e_a^{n-1} \cup e_1^n \cup \cdots \cup e_b^n, \quad \text{rel.}\, K^*.$$

Now let $k = a$. On considering the effect of a simple elementary deformation, rel. K^*, it follows inductively that (16. 1) are also satisfied by K^* and Q. Let

$$g_\rho : (E_0^n, I^{n-1}) \to (Q, \bar{e}_\rho^{n-1}) \qquad (\rho = 1, \cdots, a)$$

mean the same as in (14. 3). Since $g_\rho \mid \partial I^n = g_\rho \mid \partial E_0^n$ is homotopic in Q to a constant map, it follows from Lemma 13 that there is a formal deformation $Q \to Q'$, rel. K^n, which replaces each $S_\rho^n - e^0$ by a cell, $e_{b+\rho}^n$, with a characteristic map $g'_\rho : I^n \to \bar{e}_{b+\rho}^n$, such that $g'_\rho \mid \partial I^n = g_\rho \mid \partial I^n$. Therefore it follows, as in the proof of Lemma 15, that there is a formal deformation

$$D_2 : Q' \to Q'' = K^n \cup e'^n_1 \cup \cdots \cup e'^n_b, \quad \text{rel.}\ K^n.$$

Let $h : I^n \to \bar{e}'^n_i$ be a characteristic map for e'^n_i. Then it follows from (16. 1b) that $h_i \mid \partial I^n$ is homotopic in K^n to a constant map. Therefore it follows from Lemma 13 that there is a formal deformation

$$D_3 : Q'' \to K^n \cup S_{11}^n \cup \cdots \cup S_{1b}^n, \quad \text{rel.}\ K^n,$$

where $\{S_{1i}^n\}$ is a cluster of n-spheres attached to K^{n-1}. On reversing these constructions we have

$$P^n \cup S_{21}^n \cup \cdots \cup S_{2k}^n = D(K^n \cup S_{11}^n \cup \cdots \cup S_{1b}^n), \quad \text{rel.}\ K^n.$$

Let $P_0^n \subset P^n$ be the mapping cylinder of $\phi \mid K^{n-1}$. Then P^n is the union of P_0^n and the n-cells in K^n. Let $f : I^n \to \bar{e}^n$ be a characteristic map for an n-cell $e^n \varepsilon K^n$. Then $f \mid \partial I^n$ is obviously homotopic in P_0^n to $\phi(f \mid \partial I^n) : \partial I^n \to L^{n-1}$. Since the latter can be extended to $\phi f : I^n \to L^n$ it follows that $f \mid \partial I^n$ is homotopic in P_0^n to a constant map. Therefore

$$P^*_0 = P_0^n \cup S_{21}^n \cup \cdots \cup S_{2l}^n = D'(P^n \cup S_{21}^n \cup \cdots \cup S_{2k}^n), \quad \text{rel.}\ P_0^n,$$

where $l \geq k$ and $S_{2,k+1}, \cdots, S_{2l}^n$ are n-spheres, attached to $e'^0 \varepsilon L^n$, which correspond to the n-cells in K^n. It follows from Lemma 11 that

$$L^* = L \cup S_{21}^n \cup \cdots \cup S_{2l}^n = D^* P_0^*,$$

where D^* is a contraction. Let $\psi^* : P_0^* \to L^*$ be given by $\psi^* \mid L^* = 1$,

$\psi^*(x, t) = \phi x$ $(x \varepsilon K^{n-1})$. Then $\psi^* \varepsilon [D^*]$ and the conditions of the theorem are satisfied by a map $\psi \varepsilon [D^*D'D]$. This completes the proof.

It follows from this theorem that any two complexes of the same n-type can be interchanged by a finite sequence of elementary deformations and transformations of the form

$$(16.2) \qquad K \to L = K - e^r, \qquad L \to K = L \cup e^r \qquad (r > n),$$

where e^r is a principal cell of K. For the transformations $K \to K^n \to K^n \cup E^{n+1} \to K^n \cup S^n$ are the resultants of such sequences, where $E^{n+1} = e^0 \cup e^n \cup e^{n+1}$, $S^n = \partial E^{n+1} = e^0 \cup e^n$, and $K^n \cap E^{n+1} = e^0 \varepsilon K^0$. Conversely a formal deformation preserves the homotopy type, and hence the n-type of a complex. Also $K^n = L^n$, whence K, L are of the same n-type, if they are related by (16.2). Thus the n-type may be defined in terms of formal deformations and elementary transformations of the form (16.2).

It follows from Lemma 2 in Section **9** of CH(I) and from Lemma 13 and Theorem 12 above that, if K is any complex, there is a simplicial complex $K^* = DK$. Moreover, Sections **14-16** may be interpreted as referring to formal deformations of the kind considered in [3]. Therefore the class of simplicial complexes, which, when treated as cell-complexes, are of the same simple homotopy type, or n-type, as a given one, K, is the same as the "nucleus," or "n-group," of K, as defined in [1].

17. Homotopy systems.[40] We proceed to the simple equivalence theory of homotopy systems. In this section we confine ourselves to systems, ρ, such that $\dim \rho < \infty$ and each group ρ_n has a finite basis.

We modify the definition of a homotopy system, ρ, by associating a class of *preferred bases* with each ρ_n. Let $(a_1, \cdots, a_p)$ be a preferred basis for ρ_n. Then $(a'_1, \cdots, a'_p)$ shall be a preferred basis for ρ_n if, and only if, $a'_i = a_{j_i}^{\pm 1}$, in case $n = 1$, or $a'_i = \pm x_i a_{j_i}$ if $n > 1$, where $x_i \varepsilon \rho_1$ and $j_1, \cdots, j_p$ is a permutation of $1, \cdots, p$. We shall only admit that $f : \rho \approx \rho'$ if f carries a preferred basis for each ρ_n into a preferred basis for ρ'_n. The preferred bases for $\rho(K)$, where K is a complex, shall be the natural bases, as defined in Section **5** of CH(II). If K^0 is a single 0-cell, then the natural bases for $\rho_1(K)$ are uniquely defined. In general they depend on the choice of a tree $T \subset K^1$, which contains K^0. In this case $\rho(K)$ is the homotopy

[40] The main purpose of this section is to prove Theorem 17, which was announced in Section 7 of CH(II).

system of the pair, (K, T). However we shall continue to write it as $\rho(K)$. A complex, K, will be called a (geometrical) realization of a given system, ρ, if, and only if, $\rho \approx \rho(K)$, subject to our condition concerning preferred bases. The process of realizing ρ by K will consist of defining a particular $f : \rho \approx \rho(K)$. Having (implicitly) done this, we shall use ρ and $a \in \rho$ to denote $\rho(K)$ and fa. By a *basis* for ρ_n we shall always mean a preferred basis.

Let C and $h : \rho \rightarrow C$ be defined as in Section **8** of CH(II). Then C shall be a chain system of the kind introduced in Section **2** above, R being the group ring of $\bar{\rho}_1 = \rho_1/d\rho_2$. We insist that, if $(a_1, \cdots, a_p)$ is a (preferred) basis for ρ_n and if $(m'_1, \cdots, m'_p)$ is the basis of C_n, then $ha_i = \pm \bar{x}_i m'_{j_i}$, where $\bar{x}_i \in \bar{\rho}_1$ and $\bar{x}_i = 1$ if $n = 1$.

We are going to define a sub-system, ρ', of a homotopy system ρ. This is not quite so simple as in the case of chain systems, for the following reason. Let $\rho'_1 \subset \rho_1$ be the sub-group generated by part of a basis for ρ_1. Let $\rho'_2 \subset \rho_2$ be the sub-group generated by ρ'_1, operating on a set of elements, $(a'_1, \cdots, a'_k)$, in a basis for ρ_2. Let $d\rho'_2 \subset \rho'_1$ and let $d' : \rho'_2 \rightarrow \rho'_1$ be the homomorphism induced by $d : \rho_2 \rightarrow \rho_1$. Then ρ'_2 is not necessarily a free crossed (ρ'_1, d')-module. For example, let ρ_1 have a single free generator, x, and ρ_2 a pair of basis elements, a, b, such that $da = x$, $db = 1$. Let $\rho'_1 = \rho_1$ and let ρ'_2 be generated by ρ_1, operating on b. Since $db = 1$ we have $a + b = b + a$, whence $xb - b = a + b - a - b = 0$. Therefore ρ'_2 is not a free ρ_1-module.

Let $\bar{\rho}'_1 = \rho'_1/d\bar{\rho}'_2$ and let $i_* : \bar{\rho}'_1 \rightarrow \bar{\rho}_1$ be the homorphism induced by the identical map $i_1 : \rho'_1 \rightarrow \rho_1$.

LEMMA 16. *Let* $i_*^{-1}(1) = 1$. *Then* ρ'_2 *is a free crossed* (ρ'_1, d')*-module, having* $(a'_1, \cdots, a'_k)$ *as a basis.*

Let ρ''_2 be the free crossed (ρ'_1, d'')-module, which is defined in terms of the symbolic generators (x', α_i) and the map $\alpha_i \rightarrow da'_i$ $(i = 1, \cdots, k;$ $x' \in \rho'_1)$. Obviously $d''\rho''_2 = d\rho'_2$. Let $a''_i \in \rho''_2$ be the basis element which corresponds to the generator $(1, \alpha_i)$. It follows from Lemma 2 in Section **2** of CH(II) that an operator homomorphism, $i_2 : \rho''_2 \rightarrow \rho_2$, associated with $i_1 : \rho'_1 \rightarrow \rho_1$, is defined by $i_2 a''_i = a'_i$. Obviously $i_2\rho''_2 = \rho'_2$ and the lemma will follow when we have proved that $i_2^{-1}(0) = 0$.

Let $C_2 = h\rho_2$, $C''_2 = h''\rho''_2$ be ρ_2, ρ''_2 made Abelian and let $j : C''_2 \rightarrow C_2$ be the homomorphism induced by i_2. Since $a'_i = i_2 a''_i$ and $jh'' = hi_2$ it follows that $(jh''a''_1, \cdots, jh''a''_k)$ is part of a basis for C_2. Since $d''\rho''_2 = d\rho'_2$ and $i_*^{-1}(1) = 1$ it follows that $j^{-1}(0) = 0$. Let $a'' \in i_2^{-1}(0)$. Then

$i_1d''a'' = di_2a'' = 1$, $jh''a'' = hi_2a'' = 0$. Therefore $d''a'' = 1$, $h''a'' = 0$ and it follows from Lemma 1 in CH(II) that $a'' = 0$. Therefore $i_2^{-1}(0) = 0$ and the lemma is proved.

Let ρ'_1, ρ'_2 satisfy the conditions of Lemma 16. Let $\rho'_p \subset \rho_p$ $(p = 3, 4, \cdots)$ be the sub-group which is generated by $\bar{\rho}'_1$, operating on part of a basis for ρ_p, and let $d\rho'_p \subset \rho'_{p-1}$. Let $d' : \rho'_p \to \rho'_{p-1}$ be the homomorphism induced by $d : \rho_p \to \rho_{p-1}$. Then $\rho' = \{\rho'_p\}$, with d' as boundary operator, is a homotopy system, which we describe as a *sub-system* of ρ, on the understanding that a (preferred) basis for ρ'_n $(n \geq 1)$ is part of a basis for ρ_n.

Let ρ be a given homotopy system, let $Z_n(\rho) = d_n^{-1}(0)$ and let [41]

$$G_1(\rho) = \bar{\rho}_1, \qquad G_n(\rho) = Z_n(\rho) - d_{n+1}\rho_{n+1} \qquad (n > 1).$$

A homomorphism, $f : \rho \to \rho'$, obviously induces a family of homomorphisms $f_* : G_n(\rho) \to G_n(\rho')$ $(n = 1, 2, \cdots)$. It may be verified in the same way as in ordinary homology theory that $f_* : G_n(\rho) \approx G_n(\rho')$ if $f : \rho \equiv \rho'$. The converse is proved below.

Let $\rho' \subset \rho$ be a sub-system and let $Z_n(\rho, \rho') = d_n^{-1}\rho'_{n-1}$ $(n > 1)$. Let $a \,\varepsilon\, Z_2(\rho, \rho')$, $a' \,\varepsilon\, \rho'_2$. Then $a + a' - a = (da)a' \,\varepsilon\, \rho'_2$, since $da \,\varepsilon\, \rho'_1$. Therefore ρ'_2 is an invariant sub-group of $Z_2(\rho, \rho')$. So therefore is the direct sum $\rho'_2 + d\rho_3$, since $d\rho_3 \subset Z_2(\rho)$, which is in the centre of ρ_2. Let

$$G_n(\rho, \rho') = Z_n(\rho, \rho') - (\rho'_n + d\rho_{n+1}) \qquad (n > 1).$$

Let

$$(17.1) \qquad G_n(\rho') \xrightarrow[i_*]{} G_n(\rho) \xrightarrow[j_*]{} G_n(\rho, \rho') \xrightarrow[d_*]{} G_{n-1}(\rho') \xrightarrow[i_*]{} \cdots \xrightarrow[i_*]{} G_1(\rho)$$

be the homomorphisms, which are induced by $i : \rho' \to \rho$, the identical map $Z_n(\rho) \to Z_n(\rho, \rho')$ and by $d \mid Z_n(\rho, \rho')$. Then it may be verified, as in ordinary homology theory, that the sequence (17.1) is exact.

Let $f : \rho \to \rho'$ be a homomorphism of ρ into a system ρ', with boundary operator d'. Let $f_* : \bar{\rho}_1 \approx \bar{\rho}'_1$. We proceed to define a system, ρ^*, which we shall call the *mapping cylinder* of f. We realize the systems $\rho^2 = (\rho_1, \rho_2)$ and ρ'^2 by complexes $K = K^2$ and $K' = K'^2$, such that $K^0 = K'^0 = e^0 = K \cap K'$. By Theorem 4 in CH(II), $f : \rho^2 \to \rho'^2$ can be realized by a map $\phi : K \to K'$. Let P be the mapping cylinder of ϕ, with $e^0 \times I$ shrunk into the point e^0. Then $P^0 = e^0$. We define $\rho^*_n = \rho_n(P) = \rho_n(P^2)$ $(n = 1, 2)$.

Since ϕ induces $f_* : \bar{\rho}_1 \approx \bar{\rho}'$ and since K' is a D. R. of P, it follows that

[41] If $\rho = \rho(K)$ then $G_1(\rho) \approx \pi_1(K)$, $G_2(\rho) \approx \pi_2(K)$, $G_n(\rho) \approx H_n(K)$ if $n > 2$.

$i_*: \bar{\rho}_1 \approx \bar{\rho}^*_1$, $i'_*: \bar{\rho}'_1 \approx \bar{\rho}^*_1$, where i_*, i'_* are the injections. Therefore it follows from the proof of Lemma 16 that the injections

$$(17.2) \qquad i: \rho^2 \to \rho(P^2), \qquad i': \rho'^2 \to \rho(P^2)$$

are isomorphisms (into).

Let $\delta: \rho^2 \to \rho(P)$ be the deformation operator determined by the homotopy $\delta_t: K \to P$, which is given by $\delta_t p = (p, t)$ $(p \varepsilon K)$. Then

$$(17.3) \qquad d^*_n\delta_n = f_{n-1} - 1 - \delta_{n-1}d_n \qquad (\delta_1 d_1 = 0),$$

where $n = 2, 3$ and ρ^*_1 is written additively. Let $n \geqq 3$ and let $\delta_n\rho_{n-1}$ be a free $\bar{\rho}^*_1$-module, which is the image of ρ_{n-1} in an operator homomorphism, δ_n, whose kernel is the commutator sub-group of ρ_{n-1}. Thus $\delta_n: \rho_{n-1} \approx \delta_n\rho_{n-1}$ if $n > 3$. We take $\delta_3\delta_2 \subset \rho_3(P)$ and δ_3 shall mean the same as before. Let ρ^*_n be the direct sum $\rho^*_n = \rho'_n + \rho_n + \delta_n\rho_{n-1}$ $(n \geqq 3)$. We imbed ρ_n, ρ'_n in ρ^*_n by means of $i: \rho_n \to \rho^*_n$, $i': \rho'_n \to \rho^*_n$, where i, i' mean the same as in (17.2) if $n = 1, 2$, and $ia = (0, a, 0)$, $i'a' = (a', 0, 0)$ if $n > 2$. We define $d^*_n: \rho^*_n \to \rho^*_{n-1}$ by $d^*_n a = d_n a$, $d^*_n a' = d'_n a'$ and by (17.3), with $n \geqq 2$. If $\{a_i{}^n\}$ and $\{a'_j{}^n\}$ are bases for ρ_n and ρ'_n, then the union of $\{a_i{}^n\}$, $\{a'_j{}^n\}$ and $\{\delta_n a_k{}^{n-1}\}$ shall be a (preferred) basis for ρ^*_n. It follows from an argument in Section **8** above that $d^*_n d^*_{n+1} = 0$ if $n \geqq 3$. Also $d^*d^* = 0$ in $\rho(P)$. Therefore $d^*_n d^*_{n+1} = 0$ for every $n > 1$. Clearly d^*_n is an operator homomorphism and it follows that $\rho^* = \{\rho^*_n\}$, with $d^* = \{d^*_n\}$ as boundary operator, is a homotopy system. We call it the *mapping cylinder* of $f: \rho \to \rho'$.

Let $i': \rho' \to \rho^*$ be the identical map and let $k': \rho^* \to \rho'$ be given by

$$k'a = fa, \qquad k'a' = a', \qquad k'\delta a = 0 \qquad (a \varepsilon \rho, a' \varepsilon \rho').$$

Then $k'i' = 1$ and it is easily verified that $d'k' = k'd^*$ and that $i'k' - 1 = d^*\delta^* + \delta^*d^*$, where $\delta^*a = \delta a$, $\delta^*\rho' = \delta^*\delta\rho = 0$. Therefore $k': \rho^* \equiv \rho'$ and $k'_*: G_n(\rho^*) \approx G_n(\rho')$. Clearly $f = k'i$, where $i: \rho \to \rho^*$ is the identical map, whence $f_* = k'_*i_*$. Therefore, if each f_* is an isomorphism (onto), so is i_*. In this case it follows from the exactness of (17.1) that

$$(17.4) \qquad G_n(\rho^*, \rho) = 0 \qquad (n \geqq 1),$$

where $G_1(\rho^*, \rho) = 0$ means that $\rho^*_1 = i_*\rho_1$.

Let $r \geqq 2$ and let $(a_0{}^n, a_1{}^n, \cdots, a_{p_n}{}^n)$ be a preferred basis for ρ_n $(n = r-1, r)$. If $r = 2$ let $\rho'_1 \subset \rho_1$ be the sub-group generated by $(a_1{}^1, \cdots, a_{p_1}{}^1)$ and if $r > 2$ let $\rho'_1 = \rho_1$. If $n > 1$ $(n = r-1, r)$ let

$\rho'_n \subset \rho_n$ be the sub-group which is generated by ρ'_1 operating on $(a_1^n, \cdots, a_{p_n})$. In any case let

$$da_0^r = a_0^{r-1} - a'_0, \qquad da_i^r \in \rho_{r-1} \qquad (i = 1, \cdots, p_r),$$

where $a'_0 \in \rho'_{r-1}$ and ρ_1 is written additively if $r = 2$. Then $d\rho'_r \subset \rho'_{r-1}$ and the conditions of Lemma 16 are satisfied[42] by ρ'_1, ρ'_2. Therefore $\rho' = \{\rho'_n\}$, with $\rho'_n = \rho_n$ if $n \neq r - 1$ or r, is a sub-system of ρ. Let $i : \rho' \to \rho$ be the identical map and let $k_n : \rho_n \to \rho'_n$ the operator homomorphism, which is given by $k_n \mid \rho'_n = 1$ and $k_r a_0^r = 0$, $k_{r-1} a_0^{r-1} = a'_0$. Then it is easily verified that $kd = d'k$, whence $k : \rho \to \rho'$ is a homomorphism. We have $ki = 1$ and $ik - 1 = d\xi + \xi d$, where $\xi : \rho \to \rho$ is the deformation operator given by

$$\xi\rho' = 0, \qquad \xi a_0^r = 0, \qquad \xi a_0^{r-1} = -a_0^r.$$

Therefore $i : \rho' \equiv \rho$ and $k : \rho \equiv \rho'$. Notice that, if $f : \rho \to \rho'$ is any homomorphism such that $fi \simeq 1$, then $f \simeq fik \simeq k$.

We shall describe a homomorphism $f : \rho^0 \to \rho^1$ as an *elementary equivalence* if, and only if, ρ^0, ρ^1 are related to each other in the same way as ρ, ρ' in the preceding paragraph, and $f \cong i$ or $f \cong k$, according as $\rho^0 \subset \rho^1$ or $\rho^1 \subset \rho^0$. We shall describe a homomorphism $f : \rho \to \rho^*$ as a *simple equivalence, $f : \rho \equiv \rho_0$ (Σ)*, if, and only if, it is the resultant of a finite sequence of isomorphisms and elementary equivalences.

Let C, C' be chain systems associated with given homotopy systems ρ, ρ^1. Let $g : C \to C'$ be the chain mapping induced by a homomorphism $f : \rho \to \rho'$.

THEOREM 15. *$f : \rho \equiv \rho'(\Sigma)$ if, and only if, $g : C \equiv C'$ (Σ).*

This follows from the lemmas in Section **14** and the proof of Theorem 13, restated in terms of homotopy ssytems.

Let ρ be a homotopy system and σ a free $\bar{\rho}_1$-module, with a finite basis $(b_1, \cdots, b_q)$. Let $\rho_n^0 = \rho_n + \sigma$, $\rho_p^0 = \rho_p$ $(p \neq n)$, for a given value of $n \geq 2$. Let $d^0 : \rho_r^0 \to \rho_{r-1}^0$ be defined by $d^0 \mid \rho = d$, $d^0\sigma = 0$. Then[43] $\rho^0 = \{\rho_r^0\}$, with d^0 as boundary operator, is a homotopy system. We say that $\rho \to \rho^0$ is the result of *attaching a cluster* of n-cycles to ρ. If $\{a_i\}$ is a preferred basis for ρ_n, then $\{a_i, b_j\}$ shall be a preferred basis for ρ_n^0 and the preferred bases for ρ_p $(p \neq n)$ shall be the same in ρ^0 as in ρ.

---

[42] The homomorphism $k_1 : \rho_1 \to \rho'_1$, defined below, induces $i_*^{-1} : \rho_1 \approx \rho'_1$.

[43] If $n = 2$ then ρ_n^0 is a free crossed module since $d^0\sigma = 1$.

Let $\dim \rho$, $\dim \rho' \leq n$ $(n \geq 2)$ and let $f: \rho \to \rho'$ be a homomorphism such that

$$(17.5) \qquad f_*: G_r(\rho) \approx G_r(\rho') \qquad (r = 1, \cdots, n-1).$$

Then we have the following generalization of Tietze's theorem.

THEOREM 16. *There is a simple equivalence, $f^0: \rho^0 \equiv \rho'^0$ (Σ), such that $f^0 a = fa$ if $a \varepsilon \rho_r$ $(r < n)$, where ρ^0, ρ'^0 are formed by attaching clusters of n-cycles to ρ, ρ'.*

Since $f_*: \bar{\rho}_1 \approx \bar{\rho}'_1$ we can construct the mapping cylinder, ρ^*, of f. Then the theorem follows from the proof of Theorem 14, translated into algebraic terms.

The following corollary may be deduced from Theorem 16, or proved directly with the help of (17.4).

COROLLARY. *If $\dim \rho$, $\dim \rho' \leq n - 1$, then* (17.5) *implies $f: \rho \equiv \rho'$.*

THEOREM 17. *If $f: \rho(K) \equiv \rho'$, where K is a complex, then ρ' can be realized by a complex, K', and in such a way*[44] *that f has a realization $\phi: K \equiv K'$.*

This follows from Theorem 16 and an argument which is essentially the same as the proof of Theorem 9 on p. 1228 of [3].

18. Infinite complexes. Let K_1 be a CW-complex, as defined in CH(I), which may be infinite. Let $K_0 \subset K_1$ be a sub-complex such that $K_1 = K_0 \cup \bigcup_\alpha (e_\alpha^{n-1} \cup e_\alpha^n)$, where $\{e_\alpha^{n-1}, e_\alpha^n\}$ is an indexed aggregate of cells such that $e_\alpha^{n-1} \cup e_\alpha^n$ is an open subset of K_1 and $K_0 \to K_0 \cup e_\alpha^{n-1} \cup e_\alpha^n$ is an elementary expansion, for each α. Then $K_0 \to K_1$ will be called a *composite expansion* and $K_1 \to K_0$ a *composite contraction.* It follows from the argument used in the finite case, and (I), in Section **5** of CH(I), that K_0 is a D. R. of K_1. By a *formal deformation, $D: K \to L$,* we shall mean the resultant of a finite sequence of composite expansions and contractions. We restrict ourselves to complexes of finite dimensionality. Then the proofs of the lemmas in Section **14** and of Theorem 14 apply to infinite complexes, after a few trivial alterations.

We also admit homotopy systems of finite dimensionality, in which the

[44] In general ρ' can also be realized by a complex, K', in such a way that f has no realization $K \to K'$.

groups may have infinite bases. We define a simple equivalence, $f:\rho \equiv \rho'$, where ρ, ρ' are two such systems, by analogy with a formal deformation $D:K \to L$. Then Theorems 16, 17 can be extended without difficulty to systems in which the bases may be infinite.

It remains to be seen whether or not the purely algebraic theory developed in Section 2-9 can be extended to systems of modules with infinite bases, in such a way as to yield a generalization of Theorem 13 to infinite complexes.

MAGDALEN COLLEGE, OXFORD.

REFERENCES.

1. J. H. C. Whitehead, *Proceedings of the London Mathematical Society*, vol. 45 (1939), pp. 243-327.
2. ———, *Annals of Mathematics*, vol. 41 (1940), pp. 809-824.
3. ———, *ibid.*, vol. 42 (1941), pp. 1197-1239.
4. ———, *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 1125-32.
5. ———, *ibid.*, vol. 54 (1948), pp. 1133-45.
6. N. Jacobson, *The theory of rings*, New York (1943).
7. K. Reidemeister, *Einführung in die kombinatorische Topologie*, Brunswick (1932).
8. ———, *Journal für die reine und angewandte Mathematik*, vol. 173 (1935), pp. 164-173.
9. W. Franz, *ibid.*, vol. 176 (1937), pp. 113-134.
10. G. de Rham, *Commentarii Mathematici Helvetici*, vol. 12 (1940), pp. 191-211.
11. G. W. Higman, *Proceedings of the London Mathematical Society*, vol. 46 (1940), pp. 231-248.
12. M. H. A. Newman, *K. Akademie van Wetenschappen, Amsterdam*, vol. 29 (1926), pp. 611-626; 627-641.
13. J. W. Alexander, *Annals of Mathematics*, vol. 31 (1930), pp. 292-320.
14. S. Lefschetz and J. H. C. Whitehead, *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 510-517.
15. J. L. Kelley and Everett Pitcher, *Annals of Mathematics*, vol. 48 (1947), pp. 682-709.
16. R. H. Fox, *ibid.*, vol. 44 (1943), pp. 40-50.

LIE ALGEBRAS AND DIFFERENTIATIONS IN RINGS OF POWER SERIES.*

By G. Hochschild.

Introduction. It was proved by Ado in 1934, [1], that every Lie algebra over a field of characteristic zero can be faithfully represented by linear transformations in a finite dimensional vector space. The details of this proof, which was published in Russian, are not widely known. Ado's proof is believed to be incomplete in one point and has the further disadvantage that it is rather elaborate and does not lead to a straightforward construction of a faithful representation for a given Lie algebra.

In 1938, E. Cartan published an entirely different proof of Ado's theorem, [3], for the case where the basic field is the field of the complex numbers. In fact, Cartan proves directly the stronger result that there exists a Lie group of linear transformations whose Lie algebra is isomorphic with the given Lie algebra, and that this group can be taken to be simply connected if the Lie algebra is solvable. Cartan's proof makes use of analysis and the theory of Lie groups of transformations, but the theorem for arbitrary base fields of characteristic zero could be obtained rather easily from Cartan's theorem.

Very recently, Harish-Chandra, [7], has given a purely algebraic proof of Ado's theorem by perfecting a method which previously had been successful only in the nilpotent case (G. Birkhoff, [2]) and for 'restricted' Lie algebras of characteristic p (Jacobson, [8]).

What we shall do here is to make the analytical tools used by Cartan available to algebra by a systematic use of the theory of differential forms in rings of formal power series. More specifically, we shall apply the algebraic version of what is known as the differential calculus of Cartan [1] to show that the elements of a given Lie algebra can be represented as differentiations in a ring of power series which map a finite dimensional subspace into itself and thus yield a faithful linear representation.

By a slight modification of Cartan's procedure this program can be carried out directly for the solvable case, i. e., the preliminary study of

* Received May 2, 1948.

[1] This will be found, for instance in [4], ch. V, and in [6]. The definitions we shall give in 2 differ from those in [4] and [6] only in some inessential respects.

nilpotent Lie algebras can be short circuited. The main new difficulty arises in extending the construction to the general case. This necessitates a close study of a certain system of differential equations which is "solvable" in the sense of Lie's classical theory.

In **1** we prove a few elementary results concerning Lie algebras. These are required in the passage from a representation of the radical of a Lie algebra to a representation of the whole Lie algebra. In **2** we give an outline of the theory of differential forms which is fundamental in all the later constructions. The representation of solvable Lie algebras is dealt with in **3**, and the extension to the general case is carried out in **4**, **5**.

1. Lie algebras. We shall require a few auxiliary results concerning Lie algebras. The first of these is an easy generalization of Levi's theorem: [2]

THEOREM 1. 1. *Let P be a semisimple Lie algebra over a field K of characteristic* 0. *Suppose that H is another Lie algebra over K and that π is a homomorphism of H onto P. Then there exists an isomorphism τ of P into H such that $\pi\tau$ is the identity mapping on P.*

Proof. Let Q denote the kernel of π, R the maximal solvable ideal of H. The image under π of the sum (Q, R) is an ideal in P. Since it is isomorphic with $(Q, R)/Q$, or with $R/R \cap Q$, it is solvable. Since P is semisimple, this implies that $(Q, R)/Q = (0)$, i. e., $R \subseteq Q$.

On the other hand, H/R is semisimple and Q/R is an ideal in H/R. Hence H/R is the direct sum of Q/R and a complementary ideal which we may write S/R, where S is an ideal in H and contains R. Evidently, π maps S onto P, and the kernel of the restriction of π to S is $S \cap Q = R$. Now S/R, as a non-zero ideal of the semisimple algebra H/R, is semisimple, whence R is the maximal solvable ideal of S.

By Levi's theorem, we have a linear decomposition $S = T + R$, where T is a subalgebra which is isomorphic with S/R and therefore with P. It is now clear that there exists an isomorphism τ of P onto T which satisfies the condition of our theorem.

The next theorem is a refinement of Levi's theorem:

THEOREM 1. 2. *Let L be a Lie algebra over K, R its maximal solvable ideal. Then L is the direct sum of two ideals T and H, such that*

[2] Levi's theorem states that a Lie algebra L over a field of characteristic O can be decomposed into a linearly direct sum $S + H$, where H is the maximal solvable ideal of L and S is a semisimple subalgebra. An elegant proof is given by J. H. C. Whitehead in [9].

(1) *T is semisimple or* (0);

(2) *H contains R as its maximal solvable ideal, and if $H = P + R$ is a Levi decomposition, then no non-zero element of P effects an inner derivation in R, i.e., if $0 \neq p \in P$ there exists no r_p in R such that $r \circ p = r \circ r_p$, for every $r \in R$.*

Proof. Applying Levi's theorem we obtain a linear decomposition $L = S + R$, where S is either (0), in which case there is nothing to prove, or a semisimple subalgebra of L.

Now denote by T the set of all $s \in S$ for which there is an $r_s \in R$ with $r \circ s = r \circ r_s$, for all $r \in R$. It is easy to verify that T is an ideal in S. Hence T is either (0) or semisimple.

Let Z denote the center of R. Then, for any $t \in T$, the elements $r_t \in R$ for which $r \circ t = r \circ r_t$, for all $r \in R$, make up exactly one coset $\bar{t}$ mod Z. The mapping $t \rightarrow \bar{t}$ is easily seen to be a homomorphism of T into R/Z. Since R/Z is solvable, we must have $\bar{T} = (0)$, for otherwise it would be a semisimple subalgebra of R/Z. This means that $R \circ T = (0)$.

Since S is semisimple, it is the direct sum of T and a complementary ideal P. Hence T is a direct summand in L, and $H = P + R$ is a complementary ideal in L which satisfies condition (2) of our theorem. In fact, if one Levi decomposition of H has the property described in (2), then every Levi decomposition will have this property.

We shall later have to make reference to the maximal nilpotent ideal of a Lie algebra. In the remainder of this section we give a few facts concerning this concept.

Let us recall that a Lie algebra N is said to be nilpotent if, with $N_0 = N$, and $N_{i+1} = N_i \circ N$, there is an n such that $N_n = (0)$. As we shall see below, the sum of all nilpotent ideals of a Lie algebra is nilpotent, and it is called the maximal nilpotent ideal.

We shall take the following fundamental theorem for granted ([5], th. 3).

THEOREM 1.3. (Lie) *Let L be a solvable Lie algebra. Then every simple representation space for L is annihilated by the derived algebra $L \circ L$.*

Let M be an arbitrary representation space for the Lie algebra L. A subset S of L is said to be nilpotent on M if there exists an integer n such that, for every $m \in M$ and every set $s_1, \cdots, s_n$ of elements of S, we have $s_1 \cdots s_n \cdot m = 0$, or—as we shall indicate more briefly—if $S^n \cdot M = (0)$. It follows almost immediately from theorem 1.3 that if L is solvable then

$L \circ L$ is nilpotent on every representation space of L. In fact, we merely have to consider a composition series for the L-module M and note that $L \circ L$ maps each term of this series into the next term, because the quotients of successive terms are simple L-modules.

LEMMA 1.1. *Let M be a representation space for the Lie algebra L, S a subspace of L which is nilpotent on M. Let x be an element of L which is nilpotent on M and such that $x \circ S \subseteq S$. Then the subspace (S, x) of L which is spanned by S and x is nilpotent on M.*

Proof. There are indices p and q such that $S^p \cdot M = (0)$ and $x^q \cdot M = (0)$. Since $s \cdot (x \cdot m) = x \cdot (s \cdot m) + (x \circ s) \cdot m$, it is easy to see that every element of $(S, x)^n \cdot M$ in which altogether t operators from S occur can be written as a sum of elements each of which belongs to an $x^r \cdot (S^t \cdot M)$, with some index r. Hence the elements in which $t \geqq p$ are zero. Thus, the total number of operators from S in a non-zero element of $(S, x)^n \cdot M$ must be less than p. But then, if the total number of x's is greater than $p(q-1)$, such an element must involve x^q in at least one place, and hence is zero. Hence we must have $(S, x)^n \cdot M = (0)$ as soon as $n \geqq p(q-1) + p = pq$.

We may regard L as a representation space for any subalgebra of L, such that the transform by an element x of the element $z \in L$ is given by $x \cdot z = D_x(z) = z \circ x$. With this understanding, we have:

LEMMA 1.2. *Let R be the maximal solvable ideal of the Lie algebra L. Then $R \circ L$ is nilpotent on L.*

Proof. If x is an arbitrary element of L then (R, x) is a solvable subalgebra of L, because $(R, x) \circ (R, x) = (R \circ R, R \circ x) \subseteq R$. Hence, by the above, $(R \circ R, R \circ x)$ is nilpotent on L. Since, for every $y \in L$, $(R \circ y) \circ (R \circ R, R \circ x) \subseteq R \circ R$, it follows by repeated applications of Lemma 1.1 that $R \circ L$ is nilpotent on L.

THEOREM 1.4. *Let L be a Lie algebra, R its maximal solvable ideal. Let N be the set of all elements of R which are nilpotent on L. Then N is a nilpotent ideal of L and coincides with the sum of all nilpotent ideals of L.*

Proof. Clearly, N must contain every nilpotent ideal of L. In particular, $N \supseteq R \circ L$. If $x \in N$ it follows from Lemma 1.1 that the subspace $(x, R \circ L)$ of L is contained in N. By repeating this argument a finite number of times, noting that $y \circ (x, R \circ L) \subseteq R \circ L$, etc., we finally conclude that N is a subspace of R. It is then clear that N is actually a nilpotent ideal of L. The remaining assertion of our theorem is a trivial consequence.

COROLLARY. *The maximal nilpotent ideal of L coincides with the maximal nilpotent ideal of R.*

Proof. An element of R is nilpotent on R if and only if it is nilpotent on L.

THEOREM 1.5. *Let L, R, N be as above, and let D be a derivation in R. Then $D(R) \subseteq N$.*

Proof. We construct a Lie algebra R^* as follows: The vector space of R^* is the direct sum of R and the basic field K, i. e., the elements of R^* are pairs (x, k), where $x \varepsilon R$ and $k \varepsilon K$. We define $(x_1, k_1) \circ (x_2, k_2) = (x_1 \circ x_2 + k_2 D(x_1) - k_1 D(x_2), 0)$. Then R^* is easily seen to be a solvable Lie algebra which contains $R = (R, 0)$ as an ideal. Hence we may regard R as a representation space for R^*. Since R^* is solvable, we conclude that $R^* \circ R^* = (R \circ R, D(R))$ (subspace of R) is nilpotent on R, and hence on L. Hence $D(R) \subseteq N$.

2. Differentiations and differential forms. Let K be a field of characteristic 0. We form the ring $K\langle x_1, \cdots, x_n\rangle$ of integral (i. e., with exponents $\geqq 0$) formal power series in n variables $x_1, \cdots, x_n$ over K.

Definition 2.1. A differentiation in $K\langle x_1, \cdots, x_n\rangle$ is a mapping D of $K\langle x_1, \cdots, x_n\rangle$ into itself such that

(1) $D\{a\} = 0$, for every $a \varepsilon K$;

(2) For any two power series p, q, we have

$$D\{p + q\} = D\{p\} + D\{q\}, \text{ and } D\{pq\} = D\{p\}q + pD\{q\}.$$

Evidently, every differentiation is a K-linear transformation. Denote by $D_1, \cdots, D_n$ the partial differentiations (in the ordinary sense) with respect to $x_1, \cdots, x_n$, respectively. Then the D_i are evidently differentiations in the above sense. In fact, we have the following theorem:

THEOREM 2.1. *The set of differentiations in $K\langle x_1, \cdots, x_n\rangle$ coincides with the set of all mappings of the form $\sum_{i=1}^{n} p_i D_i$, where the p_i are arbitrary elements of $K\langle x_1, \cdots, x_n\rangle$. Hence the differentiations constitute a free $K\langle x_1, \cdots, x_n\rangle$-module of rank n.*

Proof. Let N denote the ideal of all non-units in $K\langle x_1, \cdots, x_n\rangle$.

Then N coincides with the ideal generated by $x_1, \cdots, x_n$, and $\bigcap_{e=1}^{\infty} N^e = (0)$. Clearly, if D is any differentiation, and $e > 1$, then $D(N^e) \subseteq N^{e-1}$.

Given a differentiation D, consider the differentiation

$$D' = D - \sum_{i=1}^{n} D\{x_i\} D_i.$$

It is clear that D' maps every polynomial (finite power series) into 0. Let p be an arbitrary power series. Then, given $e > 1$, we can find a polynomial p_e such that $p - p_e \,\varepsilon\, N^{e+1}$. Then $D'\{p\} = D'\{p - p_e\} \,\varepsilon\, N^e$. Hence $D'\{p\} \,\varepsilon \bigcap_{e=1}^{\infty} N^e$, i. e., $D'\{p\} = 0$, whence $D = \sum_{i=1}^{n} D\{x_i\} D_i$.

If U and V are differentiations, so is $U \circ V = VU - UV$. The set of differentiations constitutes a Lie ring with the multiplication $(U, V) \rightarrow U \circ V$.

Definition 2. 2. Let ω be a function defined on the s-fold set theoretical product of the module of differentiations by itself, and taking values in $K\langle x_1, \cdots, x_n\rangle$, which possesses the following properties:

(1) If $U_1, \cdots, U_s$ denote differentiations then $\omega\{U_1, \cdots, U_s\} = 0$, whenever two of the U_i's are equal.

(2) For fixed $U_2, \cdots, U_s$, the mapping $U \rightarrow \omega\{U, U_2, \cdots, U_s\}$ is an operator homomorphism of the $K\langle x_1, \cdots, x_n\rangle$-module of the differentiations into $K\langle x_1, \cdots, x_n\rangle$, regarded in the natural way as a module over itself.

Then ω is called a homogeneous differential form of degree s on $K\langle x_1, \cdots, x_n\rangle$. By a form of degree 0 we shall simply mean an element of $K\langle x_1, \cdots, x_n\rangle$.

Thus, speaking more loosely, a homogeneous differential form of degree s on $K\langle x_1, \cdots, x_n\rangle$ is an alternating s-linear function on the set of differentiations, taking values in $K\langle x_1, \cdots, x_n\rangle$.

It is easy to see that every homogeneous differential form of degree $> n$ must be 0. Clearly, a homogeneous differential form ω of degree s is determined completely by the values $\omega\{D_{i_1}, \cdots, D_{i_s}\}$, for $i_1 < \cdots < i_s$. Moreover, there always exists a differential form ω such that these values are arbitrarily prescribed elements of $K\langle x_1, \cdots, x_n\rangle$. It follows that, with the natural structure as a $K\langle x_1, \cdots, x_n\rangle$-module, the homogeneous differential forms of degree s constitute a free $K\langle x_1, \cdots, x_n\rangle$-module of rank $n!/s!(n-s)!$, for $0 \leqq s \leqq n$.

Next, we wish to define the Grassmann or 'outer' product of two

differential forms. To this end, we introduce the following auxiliary functions:

Let σ be the function defined on all integers such that $\sigma\{a\} = 1$, for $a > 0$; $\sigma\{a\} = -1$, for $a < 0$; $\sigma\{a\} = 0$, for $a = 0$.

If A and B are arbitrary sets of positive integers, we set

$$e(A,B) = \prod_{i\varepsilon A,\, j\varepsilon B} \sigma\{j - i\},$$

(where the product of no factors is to be interpreted as 1).

Now let θ and ω be homogeneous differential forms of degree s and t, respectively. We define a function $\theta\omega$ by setting

$$(\theta\omega)\{U_1, \cdots, U_{s+t}\} = \sum_{A,B} \theta\{U_{a_1}, \cdots, U_{a_s}\}\omega\{U_{b_1}, \cdots, U_{b_t}\},$$

where the sum is to be taken over all ordered sets of integers

$$A = (a_1, \cdots, a_s), \text{ with } 1 \leqq a_1 < \cdots < a_s \leqq s + t, \text{ and}$$

$$B = (b_1, \cdots, b_t), \text{ with } 1 \leqq b_1 < \cdots < b_t \leqq s + t.$$

Clearly, $\theta\omega$ is $(s+t)$-linear. A straightforward computation shows that $\theta\omega$ is alternating. Thus, $\theta\omega$ is a homogeneous differential form of degree $s + t$. It can be checked directly that our outer multiplication $(\theta, \omega) \to \theta\omega$ is associative and distributive. Finally, with θ and ω as above, we have $\theta\omega = (-1)^{st}\omega\theta$.

If $p \varepsilon K\langle x_1, \cdots, x_n\rangle$ we define a homogeneous differential form dp of degree 1 by setting $(dp)\{U\} = U\{p\}$, for every differentiation U. In particular, the dx_i form a set of independent generators, over $K\langle x_1, \cdots, x_n\rangle$, for the differential forms of degree 1. Their ordered products $dx_{i_1} \cdots dx_{i_s}$ constitute a set of independent generators for the homogeneous differential forms of degree s. Explicitly, we have

$$\omega = \sum_{i_1 < \cdots < i_s} \omega\{D_{i_1}, \cdots, D_{i_s}\} dx_{i_1} \cdots dx_{i_s},$$

as can be verified easily.

We wish to extend the operator d to a mapping of homogeneous differential forms of degree s into forms of degree $s + 1$, such that

(1) $dd = 0$;

(2) $d(\theta\omega) = (d\theta)\omega + (-1)^s\theta(d\omega)$, where s is the degree of θ. This is accomplished by the coboundary formula of Eilenberg and MacLane: if ω is of degree s we define

$$(d\omega)\{U_1, \cdots, U_{s+1}\} = \sum_{i=1}^{s+1} (-1)^{i-1} U_i\{\omega\{U_1, \cdots, \hat{U}_i, \cdots, U_{s+1}\}\}$$
$$+ \sum_{p<q} (-1)^{p+q-1} \omega\{U_p \circ U_q, U_1, \cdots, \hat{U}_p, \cdots, \hat{U}_q, \cdots, U_{s+1}\},$$

where the ^ indicates omission of the argument below it. The verification of the properties (1) and (2) is rather lengthy; we shall indicate only its main outline.

First, one verifies by a direct computation that (2) holds for $s = 0$, and for $s = 1$ and that $dd\theta = 0$ if θ is of degree 0. Then one establishes (2) in general by an easy induction on the degree s of θ.

Now one verifies that $d\theta$ is a differential form, i. e. is alternating and linear, by direct computation for $s = 1$ (this is trivial for $s = 0$), and induction on s thereafter.

Finally, one establishes (1), i. e., $d(d(\theta)) = 0$, by induction on the degree of θ.

The inductions are based on the use of (2) by writing a form of degree s as a sum of products of forms of degree 1 by forms of degree $s-1$.

An important fact for us is that the converse of (1) is true:

THEOREM 2. 2. *Let θ be a homogeneous differential form of degree $\geqq 1$ such that $d\theta = 0$. Then there exists a homogeneous form ϕ with $d\phi = \theta$.*

Proof. We may write θ in the form

$$\theta = \sum_{1 \leqq i_1 < \cdots < i_s \leqq n} \theta_{i_1 \ldots i_s} dx_{i_1} \cdots dx_{i_s},$$

where $\theta_{i_1 \ldots i_s} = \theta\{D_{i_1}, \cdots, D_{i_s}\}$. If ϕ is any homogeneous differential form of degree $s-1$ we write similarly $\phi = \sum \phi_{i_1 \ldots i_{s-1}} dx_{i_1} \cdots dx_{i_{s-1}}$, and we have $d\phi = \sum d\phi_{i_1 \ldots i_{s-1}} dx_{i_1} \cdots dx_{i_{s-1}} = \sum D_j\{\phi_{i_1 \ldots i_{s-1}}\} dx_j dx_{i_1} \cdots dx_{i_{s-1}}$. The condition $d\phi = \theta$ is therefore equivalent to

$$\sum_{r=1}^{s} (-1)^{r-1} D_{i_r}\{\phi_{i_1 \ldots \hat{i}_r \ldots i_s}\} = \theta_{i_1 \ldots i_s}, \qquad 1 \leqq i_1 < \cdots < i_s \leqq n.$$

We shall prove the existence of a solution by reducing the problem to the case $s = n$. (If $s > n$ we have $\theta = 0$, and the problem is trivial). If $s = n$ we have to solve only a single equation:

$$\sum_{r=1}^{n} (-1)^{r-1} D_r\{\phi_{1 \ldots \hat{r} \ldots n}\} = \theta_{1 \ldots n}.$$

This can be solved trivially by a quadrature; for instance, we may take $\phi_{1 \ldots \hat{r} \ldots n} = 0$, for $r > 1$, and determine $\phi_{2 \ldots n}$ such that $D_{11}\{\phi_{2 \ldots n}\} = \theta_{1 \ldots n}$.

If $s < n$ then, for $1 < i_2 < \cdots < i_s \leqq n$, let us find (by quadratures) elements $\phi^*_{i_2\ldots i_s}$ in $K\langle x_1, \cdots, x_n\rangle$ such that $D_1\{\phi^*_{i_2\ldots i_s}\} = \theta_{1i_2\ldots i_s}$. We now try to determine suitable elements $\phi_{i_2\ldots i_s}$, for $1 < i_2 < \cdots < i_s \leqq n$, in the form $\phi_{i_2\ldots i_s} = \phi^*_{i_2\ldots i_s} + \psi_{i_2\ldots i_s}$, where $\psi_{i_2\ldots i_s} \varepsilon K\langle x_2, \cdots, x_n\rangle$, so that $D_1\{\psi_{i_2\ldots i_s}\} = 0$. Our conditions then reduce to the following:

$$\sum_{r=2}^{s} (-1)^{r-1} D_{i_r}\{\phi_{1i_2\ldots\hat{i}_r\ldots i_s}\} = 0, \qquad \text{for } 1 < i_2 < \cdots < i_s \leqq n,$$

and

$$\sum_{r=1}^{s} (-1)^{r-1} D_{i_r}\{\phi^*_{i_1\ldots\hat{i}_r\ldots i_s} + \psi_{i_1\ldots\hat{i}_r\ldots i_s}\} = \theta_{i_1\ldots i_s}, \text{ for } 1 < i_1 < \cdots < i_s \leqq n.$$

The first set can be met simply by taking $\phi_{1i_2\ldots i_{s-1}} = 0$, $1 < i_2 < \cdots < i_{s-1} \leqq n$. The second set may be written

$$\sum_{r=1}^{s} (-1)^{r-1} D_{i_r}\{\psi_{i_1\ldots\hat{i}_r\ldots i_s}\} = \chi_{i_1\ldots i_s},$$

where

$$\chi_{i_1\ldots i_s} = \theta_{i_1\ldots i_s} + \sum_{r=1}^{s} (-1)^r D_{i_r}\{\phi^*_{i_1\ldots\hat{i}_r\ldots i_s}\}.$$

Now we have

$$\begin{aligned} D_1\{\chi_{i_1\ldots i_s}\} &= D_1\{\theta_{i_1\ldots i_s}\} + \sum_{r=1}^{s} (-1)^r D_{i_r} D_1\{\phi^*_{i_1\ldots\hat{i}_r\ldots i_s}\} \\ &= D_1\{\theta_{i_1\ldots i_s}\} + \sum_{r=1}^{s} (-1)^r D_{i_r}\{\theta_{1i_1\ldots\hat{i}_r\ldots i_s}\} \\ &= 0, \text{ since } d\theta = 0. \end{aligned}$$

Hence $\chi_{i_1\ldots i_s} \varepsilon K\langle x_2, \cdots, x_n\rangle$. Our conditions are therefore equivalent to the relation $d\psi = \chi$, in $K\langle x_2, \cdots, x_n\rangle$.

Now consider the differential form

$$\rho = \sum_{1 < i_2 < \cdots < i_s \leqq n} \phi^*_{i_2\ldots i_s} dx_{i_2} \cdots dx_{i_s}.$$

We have

$$\begin{aligned} d\rho &= \sum D_j\{\phi_{i_2\ldots i_s}\} dx_j dx_{i_2} \cdots dx_{i_s} \\ &= \sum D_1\{\phi^*_{i_2\ldots i_s}\} dx_1 dx_{i_2} \cdots dx_{i_s} \\ &\qquad + \sum_{1 < i_1 < \cdots < i_s \leqq n} (-1)^{r-1} D_{i_r}\{\phi^*_{i_1\ldots\hat{i}_r\ldots i_s}\} dx_{i_1} \cdots dx_{i_s} \\ &= \sum \theta_{1i_2\ldots i_s} dx_1 dx_{i_2} \cdots dx_{i_s} + \sum (-1)^{r-1} D_{i_r}\{\phi^*_{i_1\ldots\hat{i}_r\ldots i_s}\} dx_{i_1} \cdots dx_{i_s}. \end{aligned}$$

Hence

$$\theta - d\rho = \sum_{1 < i_1 < \cdots < i_s \leqq n} \chi_{i_1\ldots i_s} dx_{i_1} \cdots dx_{i_s}.$$

Since this involves only the x_i and dx_i for $i > 1$, its derivative as a form χ on $K\langle x_2, \cdots, x_n\rangle$ is formally the same as its derivative as a form on $K\langle x_1, \cdots, x_n\rangle$. Since $d(\theta - d\rho) = 0$, this means that χ, regarded as a differential form on $K\langle x_2, \cdots, x_n\rangle$, has derivative 0. Thus our above conditions are of the same type as the original conditions, with the number of variables reduced to $n - 1$. By repeating this reduction we finally reach the case $s = n$, and hence obtain a solution.

It is important to observe that the coefficients $\phi_{i_1 \ldots i_{s-1}}$ of the desired form ϕ can be obtained by applying quadratures and partial differentiations to the coefficients $\theta_{i_1 \ldots i_s}$.

Let L be a Lie algebra of dimension $r \leqq n$ over K. We construct the Grassmann algebra G over L whose underlying linear space is the direct sum of the spaces G_s of the s-linear alternating functions on L, with values in K, for $s = 0, 1, \cdots, r$, and where $G_0 = K$. The multiplication in G is defined exactly like the outer multiplication of differential forms. We also define a K-linear mapping δ which maps each G_s into G_{s+1}, as follows: if $g \in G_s$ $(s > 0)$ and u_i are elements of L we set

$$(\delta g)\{u_1, \cdots, u_{s+1}\} = \sum_{p<q} (-1)^{p+q-1} g\{u_p \circ u_q, u_1, \cdots, \hat{u}_p, \cdots, \hat{u}_q, \cdots, u_{s+1}\}.$$

We define $\delta(G_0) = (0)$, for completeness.

Exactly as in the case of the operator d for differential forms, we can show that $\delta(gh) = (\delta g)h + (-1)^s g(\delta h)$ and hence that $\delta g \in G_{s+1}$, and $\delta\delta = 0$.

It is no longer true here that $\delta g = 0$ implies that $g = \delta h$, with some $h \in G_{s-1}$. The additive group of the elements of G_s which are mapped into 0 by δ, modulo $\delta(G_{s-1})$ is an important invariant of L, its s-dimensional cohomology group $H^s(L)$. (See [6]).

We shall later construct an isomorphism of G into the Grassmann ring Ω of the differential forms on $K\langle x_1, \cdots, x_n\rangle$. In order that this should lead to a representation of L it is necessary that certain regularity conditions be satisfied. We shall proceed to consider this question in detail.

Denote by Ω_i the module of the homogeneous differential forms of degree i on $K\langle x_1, \cdots, x_n\rangle$. A mapping α of G into Ω will be called an isomorphism if the following conditions are satisfied:

(1) On $G_0 = K$, α is the natural injection of K into $K\langle x_1, \cdots, x_n\rangle$.

(2) For each i, α maps G_i isomorphically into Ω_i.

(3) $\alpha(gg') = \alpha(g)\alpha(g')$, and $\alpha(\delta g) = d\alpha(g)$, for all $g, g' \in G$.

Since, for $s > 0$, $G_s = (G_1)^s$, it follows from (3) that an isomorphism α is completely determined by its restriction to G_1. The regularity condition which we wish to impose is the following: denote by Q_r the submodule of Ω_r consisting of all forms ω for which $\omega\{U_1, \cdots, U_r\} \varepsilon N$, the ideal of non-units, for all differentiations U_i. Then we shall say that the isomorphism α is regular if $\alpha(G_1)^r \not\subseteq Q_r$.

Let $g_1, \cdots, g_r$ be any basis for G_1 over K. Then $\alpha(G_1)^r$ consists of all the K-multiples of the single element $\alpha(g_1) \cdots \alpha(g_r)$, and our regularity condition is equivalent to the condition $\alpha(g_1) \cdots \alpha(g_r) \not\varepsilon Q_r$. Each $\alpha(g_i)$ may be written in the form $\alpha(g_i) = \sum_{j=1}^{n} a_{ij} dx_j$, where the $a_{ij} \varepsilon K\langle x_1, \cdots, x_n \rangle$, and we have $\alpha(g_1) \cdots \alpha(g_r) = \sum | a_{ij} |_{j\varepsilon A}\, dx_{\alpha_1} \cdots dx_{\alpha_r}$, where the sum is taken over all sets A: $1 \leqq \alpha_1 < \cdots < \alpha_r \leqq n$. Hence, if α is regular, there exists at least one such set A for which the determinant $| a_{ij} |_{j\varepsilon A} \not\varepsilon N$. If $B = (\beta_1, \cdots, \beta_{n-r})$ is the complementary set of indices this evidently gives $dx_{\alpha_i} = \sum_{j=1}^{r} b_{ij} \alpha(g_j) + \sum_{k=1}^{n-r} c_{ik} dx_{\beta_k}$, with $b_{ij}, c_{ik} \varepsilon K\langle x_1, \cdots, x_n \rangle$. Hence the elements $\alpha(g_1), \cdots, \alpha(g_r)$; dx_{β_k} constitute a system of generators for Ω_1. This system is independent since

$$\alpha(g_1) \cdots \alpha(g_r) dx_{\beta_1} \cdots dx_{\beta_{n-r}} = e(A, B) \, | a_{ij} |_{j\varepsilon A}\, dx_1 \cdots dx_n \neq 0.$$

We are now in a position to prove the following theorem:

THEOREM 2.3. *Let L be a Lie algebra of dimension r over K, and let G be the associated Grassmann algebra. Suppose there is given a regular isomorphism $g \to \bar{g}$ of G into the module Ω of the differential forms on $K\langle x_1, \cdots, x_n \rangle$. Then there is an isomorphism $u \to \bar{u}$ of L onto a Lie algebra of differentiations in $K\langle x_1, \cdots, x_n \rangle$, such that $g\{u\} = \bar{g}\{\bar{u}\}$ for every $g \varepsilon G$ and $u \varepsilon L$.*

Proof. Let $g_1, \cdots, g_r$ be a basis for G_1 over K. By the above, we may suppose that $\bar{g}_1, \cdots, \bar{g}_r, dx_{r+1}, \cdots, dx_n$ form a free system of generators for Ω_1 over $K\langle x_1, \cdots, x_n \rangle$. Then it is clear that, for each $u \varepsilon L$, there exists a unique differentiation $\bar{u}$ in $K\langle x_1, \cdots, x_n \rangle$, such that $(dx_j)\{\bar{u}\} = 0$ for $j > r$, and $\bar{g}_i\{\bar{u}\} = g_i\{u\}$.

We shall prove that the mapping $u \to \bar{u}$ is an isomorphism. Since it is evidently a K-linear isomorphism of the vector space L into the module of differentiations, there remains only to show that $\overline{u \circ v} = \bar{u} \circ \bar{v}$. Now for $g \varepsilon G_1$ we have $(d\bar{g})\{\bar{u}, \bar{v}\} = \bar{u}\{\bar{g}\{\bar{v}\}\} - \bar{v}\{\bar{g}\{\bar{u}\}\} + \bar{g}\{\bar{u} \circ \bar{v}\} = \bar{g}\{\bar{u} \circ \bar{v}\}$, since $\bar{g}\{\bar{u}\} = g\{u\} \varepsilon K$, and $\bar{g}\{\bar{v}\} = g\{v\} \varepsilon K$. Write $\delta g = \sum_{p<q} a_{pq} g_p g_q$, where $a_{pq} \varepsilon K$.

Then

$$(\delta g)\{u, v\} = \sum_{p<q} a_{pq}(g_p\{u\}g_q\{v\} - g_p\{v\}g_q\{u\})$$

$$= \sum_{p<q} a_{pq}(\bar{g}_p\{\bar{u}\}\bar{g}_q\{\bar{v}\} - \bar{g}_p\{\bar{v}\}\bar{g}_q\{\bar{u}\}) = \overline{\delta g}\{\bar{u}, \bar{v}\} = (d\bar{g})\{\bar{u}, \bar{v}\}$$

i.e., $g\{u \circ v\} = \bar{g}\{\bar{u} \circ \bar{v}\}$, i.e., $\bar{g}\{\overline{u \circ v}\} = \bar{g}\{\bar{u} \circ \bar{v}\}$, which shows that $\overline{u \circ v} = \bar{u} \circ \bar{v}$.

3. Representation of solvable Lie algebras. We shall base our computations on the following known theorem:

THEOREM 3.1. *Let L be a solvable Lie algebra over a field F of characteristic 0. Then there exists a finite algebraic extension field K of F such that the extension L^0 of L over K has a composition series of the following sort: $L^0 = L_0{}^0 \supset L_1{}^0 \supset \cdots \supset L_n{}^0 = (0)$, where $L_i{}^0$ is an ideal in L^0 and is of dimension $n - i$ over K. Moreover, there is an index r such that $L_r{}^0$ coincides with the maximal nilpotent ideal of L^0.*

Proof. Let F_1 be the algebraic closure of F. Denote by L^1 the extension of L over F_1, and by N^1 the extension over F_1 of the maximal nilpotent ideal N of L. Since N^1 is an ideal in L^1, we may regard it as an L^1-module in the natural fashion. Let $N^1 = Q_0 \supset Q_1 \supset \cdots \supset Q_s = (0)$ be any composition series of this module. Then the quotients Q_{i-1}/Q_i are simple representation spaces for L^1. By theorem 1.3, every element of L^1 induces a transformation in Q_{i-1}/Q_i which commutes with every other transformation belonging to our representation. Making use of Schur's lemma and the fact that F_1 is algebraically closed, we conclude that all these transformations are scalar multiplications. Since Q_{i-1}/Q_i is simple, this implies that it is of dimension 1 over F_1. This means that the ideals Q_i are of dimension $s - i$ over F_1. We choose elements $\nu_1, \cdots, \nu_s$ such that $\nu_i \in Q_{i-1}$ but $\nu_i \notin Q_i$. Then this set constitutes a basis for N^1 over F_1. Each ν_i may be expressed in terms of a basis for L over F, allowing the coefficients to lie in F_1. This finite set of coefficients generates a finite algebraic extension K of F.

Let N^0 be the extension of N over K. If $L^0{}_{n-i}$ is the vector space over K which is spanned by $\nu_s, \nu_{s-1}, \cdots, \nu_{s-i+1}$ then these $L^0{}_j$ are ideals in L^0 and $L^0{}_{n-s} = N^0$. The maximal nilpotent ideal of L^0 must evidently contain N^0. Since $L^0 \circ L^0 = (L \circ L)^0 \subset N^0$, every subspace of L^0 which contains N^0 is an ideal, and we can trivially determine $L_1{}^0, \cdots, L^0{}_{n-s-1}$ so as to satisfy the conditions of Theorem 3.

Henceforth we shall suppose that L is a solvable Lie algebra over K which already possesses a composition series

$$L = L_0 \supset L_1 \supset \cdots \supset L_r = N \supset L_{r+1} \supset \cdots \supset L_n = (0),$$

where each L_i is an ideal in L and is of dimension $n - i$ over K. We select a basis $u_1, \cdots, u_n$ of L over K such that $u_i \varepsilon L_{i-1}$, $u_i \not\varepsilon L_i$. If we write $u_i \circ u_j = \sum_{k=1}^{n} c_{ijk} u_k$, the $c_{ijk} \varepsilon K$ have the following properties:

(1) $c_{ijk} = 0$, unless $k \geqq i$, $k \geqq j$, and $k > r$.

(2) If $i > r$ and $j > r$ then $c_{ijk} = 0$, unless $k > i$ and $k > j$.

In fact, (1) follows from the construction of the u_i and the fact that $L \circ L \subseteq N$, the maximal nilpotent ideal. (2) follows from the fact that $u_i \varepsilon N$ for $i > r$ and that N is nilpotent.

Let $g_1, \cdots, g_n$ be the basis for the elements of degree 1 in the Grassmann algebra over L for which $g_i\{u_j\} = \delta_{ij}$. Then we have $(\delta g_k)\{u_i, u_j\} = g_k\{u_i \circ u_j\} = c_{ijk}$, whence, using (1), $\delta g_k = \sum_{i<j\leqq k} c_{ijk} g_i g_j$.

We wish to construct a regular isomorphism $g \to \bar{g}$ of the Grassmann algebra G over L into Ω, the module of the homogeneous differential forms on $K\langle x_1, \cdots, x_n\rangle$. We shall do this by constructing successively $\bar{g}_1 = \omega_1, \cdots, \bar{g}_n = \omega_n$ so as to satisfy all the conditions laid down in **2**.

First, we note that, for $k \leqq r$, we have $\delta g_k = 0$. Therefore, we may define $\omega_k = dx_k$, for $k \leqq r$.

We shall prove that we can find forms ω_k such that $d\omega_k = \sum_{i<j\leqq k} c_{ijk} \omega_i \omega_j$ and that, moreover, they may be taken in the form

$$\omega_k = f_k dx_k + \sum_{s=1}^{k-1} P_{ks}(f_1, \cdots, f_{k-1}; x_1, \cdots, x_{k-1}) dx_s,$$

where the P_{ks} are polynomials, and where $f_i = \exp\left(\sum_{p=1}^{r} c_{pii} x_p\right)$. Note that, by (1), $f_i = 1$ for $i \leqq r$, so that our assertion holds true for $\omega_1, \cdots, \omega_r$.

Now let $k > r$, and suppose that $\omega_1, \cdots, \omega_{k-1}$ have already been determined. The above differential equation for ω_k may be written:

$$d\omega_k - \left(\sum_{i=1}^{k-1} c_{ikk} \omega_i\right) \omega_k = \sum_{i<j<k} c_{ijk} \omega_i \omega_j.$$

By the property (2) of the c_{ijk} we see that the sum on the left has

non-zero terms only up to $i = r$. Hence the ω_i on the left are the dx_i. If we multiply by $f^{-1}{}_k$ the equation takes the form

$$d(f^{-1}{}_k\omega_k) = f^{-1}{}_k \sum_{i<j<k} c_{ijk}\omega_i\omega_j = \sigma_k,$$

say.

We wish to verify that $d\sigma_k = 0$. Direct computation gives

$$d\sigma_k = f^{-1}{}_k[d(\sum_{i<j<k} c_{ijk}\omega_i\omega_j) - \sum_{p=1}^{r} c_{pkk}\omega_p \sum_{i<j<k} c_{ijk}\omega_i\omega_j].$$

On the other hand, since $\delta(\delta g_k) = 0$, we have

$$\begin{aligned}\delta(\sum_{i<j<k} c_{ijk}g_ig_j) &= -\delta(\sum_{p=1}^{r} c_{pkk}g_pg_k)\\ &= -\sum_{p=1}^{r} c_{pkk}(\delta g_p)g_k + \sum_{p=1}^{r} c_{pkk}g_p\delta g_k\\ &= \sum_{p=1}^{r} c_{pkk}(\delta g_p)g_k + \sum_{p=1}^{r} c_{pkk}g_p(\sum_{i=1}^{r} c_{ikk}g_ig_k + \sum_{i<j<k} c_{ijk}g_ig_j).\end{aligned}$$

Noting that $\delta g_p = 0$ for $p \leqq r$, and that the square of a homogeneous element of degree 1 in a Grassmann algebra is 0, we see that this relation reduces to

$$\delta(\sum_{i<j<k} c_{ijk}g_ig_j) = \sum_{p=1}^{r} c_{pkk}g_p \sum_{i<j<k} c_{ijk}g_ig_j.$$

Since, by our inductive hypothesis, the ω_i must satisfy the same relations over K as the g_i for $i = 1, \cdots, k-1$, the last relation implies that $d\sigma_k = 0$.

Now we may apply Theorem 2.2 and conclude that there exists a homogeneous differential form ϕ_k of degree 1 such that $d\phi_k = \sigma_k$. Since the coefficients of ϕ_k can be obtained by applying partial differentiations and quadratures to the coefficients of σ_k, it follows from our inductive assumption about the form of the ω_i for $i < k$ that ϕ_k may be taken of the form

$\phi_k = f^{-1}{}_k \sum_{s=1}^{k-1} P_{ks}(f_1, \cdots, f_{k-1}; x_1, \cdots, x_{k-1})\,dx_s$, where the P_{ks} are polynomials. We set $\omega_k = f_k(dx_k + \phi_k)$, i.e.,

$$\omega_k = f_kdx_k + \sum_{s=1}^{k-1} P_{ks}(f_1, \cdots, f_{k-1}; x_1, \cdots, x_{k-1})\,dx_s.$$

Then it can be verified quite easily that $d\omega_k = \sum_{i<j\leqq k} c_{ijk}\omega_i\omega_j$.

Hence we can determine $\omega_1, \cdots, \omega_n$ so as to satisfy the above conditions. Furthermore, we have then

$$\omega_1 \cdots \omega_n = \prod_{k=1}^{n} (f_kdx_k + \sum_{s=1}^{k-1} P_{ks}dx_s) = (\prod_{k=1}^{n} f_k)\,dx_1 \cdots dx_n.$$

Since $\prod_{k=1}^{n} f_k = \exp\,(\sum c_{pkk}x_p)$ is a unit, we see that the correspondence $g_i \to \omega_i$ can be extended to a regular isomorphism of G into Ω. By Theorem 2. 3, there is an isomorphism $u \to \bar{u}$ of L onto a Lie algebra of differentiations in $K\langle x_1, \cdots, x_n\rangle$ such that $g_i\{u\} = \omega_i\{\bar{u}\}$. From the form of the ω_i we conclude that

(1) $\bar{u}\{x_i\} \,\varepsilon\, K$, for $i \leqq r$;

(2) $\bar{u}\{x_j\}$ is a polynomial in $f_1, \cdots, f_j;\ f^{-1}{}_1, \cdots, f^{-1}{}_j;\ x_1, \cdots, x_{j-1}$;

(3) If $j > r$ and $i \leqq r$ then $\bar{u}_j\{x_i\} = 0$.

It follows that

(4) $\bar{u}\{f_k\} = af_k$, with $a \,\varepsilon\, K$;

(5) If $j > r$ then $\bar{u}_j\{f_k\} = 0$.

We wish to show next that repeated application of the $\bar{u}$ to the x_j generates only a finite dimensional K-linear subspace of $K\langle x_1, \cdots, x_n\rangle$. It is clear that every element of $K\langle x_1, \cdots, x_n\rangle$ which can be obtained from the x_i by repeatedly applying operators $\bar{u}$ may be written in the form $\sum_\alpha f_1{}^{e_1(\alpha)} \cdots f_n{}^{e_n(\alpha)} P_\alpha(x_1, \cdots, x_n)$, where the P_α are polynomials and the $e_i(\alpha)$ are positive or negative integers. There are at most n^2 linearly independent first transforms $\bar{u}\{x_i\}$. Let $M > 1$ be an upper bound for the total degrees of the P_α which occur in these transforms. We introduce a weight function w for polynomials $P(x_1, \cdots, x_n)$ as follows:

If m_i is the degree of $P(x_1, \cdots, x_n)$ in x_i we define $w(P) = \sum_{i=1}^{n} m_i M^{2i}$. Then, if P is of degree $m_i \neq 0$ in x_i, we have $D_i\{P\}$ of degree $m_i - 1$ in x_i. This shows that $\bar{u}\{P\} = \sum_{i=1}^{n} D_i\{P\}\bar{u}\{x_i\}$ can be written in the above form with all the polynomials P_α of weight no greater than

$$\operatorname{Max}_i[w(P) - M^{2i} + M \cdot M^{2(i-1)}] < w(P).$$

Now we have

$$\bar{u}\{\sum_\alpha f_1{}^{e_1(\alpha)} \cdots f_n{}^{e_n(\alpha)} P_\alpha\} = \sum_\alpha f_1{}^{e_1(\alpha)} \cdots f_n{}^{e_n(\alpha)} a(\alpha) P_\alpha + \sum_\alpha f_1{}^{e_1(\alpha)} \cdots f_n{}^{e_n(\alpha)} \bar{u}\{P_\alpha\}$$

where the $a(\alpha) \,\varepsilon\, K$. If the second sum is written in the standard form, the exponents $e_i(\alpha)$ may be increased numerically, but the weights of all the new polynomials which appear are less than the maximum of the $w(P_\alpha)$.

It follows that all the repeated transforms of the x_i may be written in the standard form in such a way that there is an upper bound for all the $|e_i(\alpha)|$ which appear, and such that each polynomial P_α is of weight no greater than $w(x_n) = M^{2n}$. Evidently, all these elements are contained in a finite dimensional K-linear subspace of $K\langle x_1, \cdots, x_n\rangle$. Furthermore, it follows from the above and from (5) that $\bar{N}$ consists entirely of nilpotent differentiations. We may state our result as follows:

THEOREM 3.2. *Let L be a solvable Lie algebra over a field F of characteristic 0. Then there exists a finite algebraic extension K of F and an isomorphism $u \to \bar{u}$ of L onto a Lie algebra $\bar{L}$ of differentiations in $K\langle x_1, \cdots, x_n\rangle$, where n is the dimension of L over F. There is a finite dimensional F-linear subspace U of $K\langle x_1, \cdots, x_n\rangle$ which contains the x_i and is such that $\bar{L}\{U\} \subseteq U$. If $\tilde{u}$ is the restriction to U of $\bar{u}$, the mapping $u \to \tilde{u}$ is an isomorphism of L onto a Lie algebra $\tilde{L}$ of linear transformations in U. Moreover, if N is the maximal nilpotent ideal of L then $\bar{N}$ consists entirely of nilpotent differentiations, and hence $\tilde{N}$ consists entirely of nilpotent transformations.*

4. Adjunction of derivations. Let L be a solvable Lie algebra and suppose that the elements of L are identified with the differentiations in $K\langle x_1, \cdots, x_n\rangle$ which were obtained in **3**. Let τ be any derivation in L and write $\tau\{u_i\} = \sum_{j=1}^{n} a_{ij}u_j$. We wish to find a differentiation t in $K\langle x_1, \cdots, x_n\rangle$ such that $tu - ut = \tau\{u\}$ for every $u \varepsilon L$.

We may assume that t is of the form $t = \sum_{i=1}^{n} g_i u_i$, with $g_i \varepsilon K\langle x_1, \cdots, x_n\rangle$, and we have to determine the g_i so as to satisfy the relations $tu_j - u_j t = \sum_{k=1}^{n} a_{jk}u_k$, i. e.,

$$\sum_{i=1}^{n} g_i(u_i u_j - u_j u_i) - \sum_{i=1}^{n} u_j\{g_i\}u_i = \sum_{k=1}^{n} a_{jk}u_k,$$

i. e., $-\sum_{i=1}^{n} g_i c_{ijk} - u_j\{g_k\} = a_{jk}$ or, using the property (1), **3**, of the c_{ijk},

$$(1) \qquad u_j\{g_k\} = -\left(a_{jk} + \sum_{i=1}^{k} g_i c_{ijk}\right).$$

Let $\omega_1, \cdots, \omega_n$ be the differential forms of degree 1 on $K\langle x_1, \cdots, x_n\rangle$ for which $\omega_i\{u_j\} = \delta_{ij}$. Then the last relations give

$$dg_k = -\sum_{j=1}^{n} (a_{jk} + \sum_{i=1}^{k} g_i c_{ijk}) \omega_j.$$

This may be written

$$dg_k - \sum_{j=1}^{r} g_k c_{jkk} \omega_j = -\sum_{j=1}^{n} (a_{jk} + \sum_{i<k} g_i c_{ijk}) \omega_j.$$

Since $\omega_j = dx_j$ for $j \leqq r$, this gives

$$(2) \qquad dh_k = -f^{-1}{}_k \sum_{j=1}^{n} (a_{jk} + \sum_{i<k} g_i c_{ijk}) \omega_j$$

where $h_k = f^{-1}{}_k g_k$. For $k \leqq r$ our equations reduce to $dh_k = -\sum_{j=1}^{n} a_{jk} \omega_j$, i. e., $dg_k = -\sum_{j=1}^{n} a_{jk} \omega_j$. By Theorem 1. 5, $\tau\{N\} \subseteq N$, whence $a_{jk} = 0$ if $j > r$ and $k \leqq r$. Therefore, the last equations reduce to $dg_k = -\sum_{j=1}^{r} a_{jk} dx_j$, which can be satisfied with $g_k = -\sum_{j=1}^{r} a_{jk} x_j$, for $k \leqq r$.

Suppose that we have already found $g_1, \cdots, g_{s-1}$ (and hence $h_1, \cdots, h_{s-1}$) such that (1) (and therefore (2)) hold for all $k \leqq s-1$. Set

$$\rho_s = f^{-1}{}_s \sum_{j=1}^{n} (a_{js} + \sum_{i<s} g_i c_{ijs}) \omega_j.$$

We wish to show that $d\rho_s = 0$. We shall do this by proving that, for all p, q, $u_q\{\rho_s\{u_p\}\} - u_p\{\rho_s\{u_q\}\} = \rho_s\{u_p \circ u_q\}$.

Now $\rho_s\{u_p\} = f^{-1}{}_s(a_{ps} + \sum_{i<s} g_i c_{ips})$. A straightforward computation using (1) and the properties of the c_{ijk} gives

$$f_s(u_q\{\rho_s\{u_p\}\} - u_p\{\rho_s\{u_q\}\}) = \sum_{i=1}^{n} (a_{pi} c_{iqs} - a_{qi} c_{ips}) + \sum_{j=1}^{s-1} g_j \sum_{i=1}^{n} (c_{jpi} c_{iqs} - c_{jqi} c_{ips}).$$

Since

$$\tau\{u_p\} \circ u_q - \tau\{u_q\} \circ u_p = \tau\{u_p \circ u_q\} = \sum_{i=1}^{n} c_{pqi} \tau\{u_1\},$$

we have

$$\sum_{i=1}^{n} (a_{pi} c_{iqs} - a_{qi} c_{ips}) = \sum_{i=1}^{n} c_{pqi} a_{is}.$$

Since

$$(u_j \circ u_p) \circ u_q - (u_j \circ u_q) \circ u_p = u_j \circ (u_p \circ u_q),$$

we have

$$\sum_{i=1}^{n} (c_{jpi}c_{iqs} - c_{jqi}c_{ips}) = \sum_{i=1}^{n} c_{pqi}c_{jis}.$$

Hence

$$f_s(u_q\{\rho_s\{u_p\}\} - u_p\{\rho_s\{u_q\}\}) = \sum_{i=1}^{n} c_{pqi}(a_{is} + \sum_{j<s} g_j c_{jis}) = f_s\rho_s\{u_p \circ u_q\},$$

which proves that $d\rho_s = 0$.

By Theorem 2. 2, there exists an element $h_s \varepsilon K\langle x_1, \cdots, x_n\rangle$ such that $dh_s = -\rho_s$. If we put $g_s = f_s h_s$ we see that g_s satisfies (1). Thus, the required differentiation exists.

Next we wish to investigate the effect of applying t to the elements of the space U of Theorem 3. 2. An important fact is that $t\{f_k\} = 0$, for all k. We have $t\{f_k\} = f_k \sum_{p=1} t\{x_p\} c_{pkk}$. Since, for $p \leqq r$, $t\{x_p\} = g_p = -\sum_{j=1}^{r} a_{jp}x_j$, we obtain

$t\{f_k\} = -f_k \sum_{j=1}^{r} (\sum_{p=1}^{r} a_{jp}c_{pkk})x_j$. We shall show that $\sum_{p=1}^{r} a_{jp}c_{pkk} = 0$.

If $u = \sum_{i=1}^{n} e_i u_i$ we have

$$u_k \circ u = \sum_{i=1}^{n} e_i \sum_{q=1}^{n} c_{kiq}u_q = -\sum_{i=1}^{n} e_i c_{ikk} u_k - \sum_{q>k} (\sum_{i=1}^{n} e_i c_{ikq}) u_q,$$

which shows that $-\sum_{i=1}^{n} e_i c_{ikk}$ is a characteristic root of D_u for every k. Hence if $u \varepsilon N$ then $\sum_{i=1}^{n} e_i c_{ikk} = 0$. By Theorem 1. 5., $\tau(u_j) \varepsilon N$. We may therefore conclude that $\sum_{r=1}^{r} a_{jp}c_{pkk} = 0$, whence $t\{f_k\} = 0$.

We have seen in **3** that the elements of U can be written

$$\phi = \sum_{a} f_1^{e_1(a)} \cdots f_n^{e_n(a)} P_a(x_1, \cdots, x_n).$$

Hence

$$t\{\phi\} = \sum_{a} f_1^{e_1(a)} \cdots f_n^{e_n(a)} \sum_{i=1}^{n} g_i u_i\{P_a\}.$$

The elements $u_i\{P_a\}$ can be written as sums of products of monomials in the f_j (allowing negative exponents) by polynomials in the x_k of lower weight than P_a.

For greater convenience, we introduce the following notion: if ϕ is any element of the above form, not necessarily contained in U, we shall define the weight of ϕ as the minimum, for all standard representations, of the greatest occurring $w(P_a)$. Any polynomial in the f_i, f^{-1}_i and x_i will henceforth be called a standard element.

Now we may express the result we have just obtained by saying that if ϕ is any standard element then $t\{\phi\} = \sum_{i=1}^{n} g_i\phi_i$, where the ϕ_i are standard elements of lower weight than ϕ.

Let us observe further that, by (1), we have $t\{g_k\} = -\sum_{i=1}^{n} g_i a_{ik}$. If $\phi = \sum_\gamma G_\gamma H_\gamma$, where the G_γ are polynomials in the g_i and the H_γ are standard elements, we have $t\{\phi\} = \sum_\gamma t\{G_\gamma\}H_\gamma + \sum_{\gamma,i}(G_\gamma g_i)H'_{\gamma i}$. The degree of $t\{G_\gamma\}$ is no greater than the degree of G_γ and the weight of each $H'_{\gamma i}$ is less than the weight of H_γ. Since t maps elements of weight 0 into 0, we conclude that all the repeated transforms of a standard element by t lie in a finite-dimensional K-linear subspace of $K\langle x_1, \cdots, x_n\rangle$. It follows that there exists a finite dimensional K-linear subspace $U_1 \supseteq U$ of $K\langle x\ , \cdots, x_n\rangle$ which is mapped into itself by t, and which is spanned by repeated t-transforms of elements of U.

Since $ut = tu - \tau\{u\}$, it is easily seen that the K-linear subspace spanned by repeated transforms of the elements of U by all the operations from (L, t) coincides with the K-linear subspace spanned by the repeated t-transforms of the elements of U, i. e., with U_1.

Finally, we note that it follows inductively from (1) that all the g_k are standard elements, and hence that U_1 consists entirely of standard elements. We may state our result as follows:

THEOREM 4.1. *Let L be a Lie algebra over K whose elements are differentiations in $K\langle x_1, \cdots, x_n\rangle$ with the properties listed in* **3**. *Let τ be an arbitrary derivation of L. Then there exists a differentiation t in $K\langle x_1, \cdots, x_n\rangle$ such that $tu - ut = \tau\{u\}$, for all $u \in L$. Furthermore, if V is any finite dimensional K-linear subspace of $K\langle x_1, \cdots, x_n\rangle$ which consists entirely of standard elements, there exists a finite dimensional subspace $V_1 \supseteq V$ of $K\langle x_1, \cdots, x_n\rangle$ such that all the operations from (L, t) map V_1 into itself, and V_1 consists entirely of standard elements.*

Only the last statement requires comment: we merely note that the proof of Theorem 3.2 shows that the operations from L generate from V

only a finite dimensional subspace consisting of standard elements. The argument given above now suffices to establish Theorem 4.1.

A differentiation t in $K\langle x_1, \cdots, x_n\rangle$ will be called singular if (1) $tu - ut = 0$, for all $u \varepsilon L$, and (2) $t\{x_i\} = 0$ for $i \leqq r$.

We wish to show that the set of singular differentiations is of finite dimension over K. For a singular differentiation $t = \sum_{i=1}^{n} g_i u_i$ the differential equations (2) become

$$dh_k = -f^{-1}{}_k \sum_{j=1}^{n} \sum_{i<k} g_i c_{ijk} \omega_j.$$

If $g_1, \cdots, g_s = 0$ these equations are satisfied for $k \leqq s$. We must have $dh_{s+1} = 0$, whence $g_{s+1} = a_{s+1} f_{s+1}$; $a_{s+1} \varepsilon K$. It follows from our proof of Theorem 4.1 that for each $i > r$ we can find a singular differentiation t_i of the form $t_i = f_i u_i + \sum_{j>i} g_{ij} u_j$.

If $s \geqq r$ the differentiation $t - a_{s+1} t_{s+1}$ is singular, and if $t - a_{s+1} t_{s+1} = \sum g'_i u_i$ we have $g'_i = 0$, for all $i \leqq s + 1$. It follows that every singular differentiation is a K-linear combination of $t_{r+1}, \cdots, t_n$.

5. Representation of arbitrary Lie algebras. Let H be an arbitrary Lie algebra over the field F of characteristic 0. Denote the maximal solvable ideal of H by L. By Theorem 1.2, we have a linearly direct decomposition $H = T + P + L$, where T is (0) or a semisimple ideal of H such that $T \circ (P + L) = (0)$, and where P is (0) or a semisimple subalgebra of H such that if $0 \neq p \varepsilon P$ the derivation $u \rightarrow u \circ p$ is not an inner derivation in L. Since a semisimple Lie algebra has, trivially, a faithful linear representation (e.g. its adjoint representation), we may confine ourselves to the case where $T = (0)$, i.e. $H = P + L$, and there remains only the case where $P \neq (0)$.

Let $p_1, \cdots, p_m$ be a basis for P over F. By Theorem 3.2, we may suppose that L is identified with a Lie algebra of differentiations in $K\langle x_1, \cdots, x_n\rangle$ as in **3**, **4**, where K is a finite algebraic extension of F. By Theorem 4.1, we can find differentiations t_i in $K\langle x_1, \cdots, x_n\rangle$ such that $t_i u - u t_i = u \circ p_i$ for every $u \varepsilon L$. For $p = \sum_{i=1}^{m} a_i p_i$, $a_i \varepsilon K$, we set $p^* = \sum_{i=1}^{m} a_i t_i$, so that $p^*{}_i = t_i$.

Now we claim that the p^* for $p \varepsilon P$ generate a finite dimensional Lie algebra over K (and hence also a finite dimensional Lie algebra over F) with the commutation $(p^*, q^*) \rightarrow p^* \circ q^* = q^* p^* - p^* q^*$. In order to prove this, we note first that the differentiations $p^* \circ q^* - (p \circ q)^*$, for $p, q \varepsilon P$, are

singular. In fact, from the relations $p^*u - up^* = u \circ p$ it follows at once that each $p^* \circ q^* - (p \circ q)^*$ commutes with every $u \varepsilon L$. Furthermore, for $k \leqq r$, we have $p^*\{x_k\} = -\sum_{s=1}^{r} a_{sk}x_s$, where the coefficients $a_{sk} \varepsilon K$ are determined from the defining relations $u_s \circ p = \sum_{k=1}^{n} a_{sk}u_k$ in L. The derivation $u \to u \circ p$ induces in L/N (N the maximal nilpotent ideal) a derivation $\bar{p}$ such that—if $\bar{u}$ denotes the coset mod N of $u \varepsilon L$—we have $\bar{p}\{\bar{u}\} = \sum_{k=1}^{r} a_{sk}\bar{u}_k$. Let p, q be elements of P and denote the matrices of $\bar{p}$, $\bar{q}$ with respect to the basis $\bar{u}_1, \cdots, \bar{u}_r$ of L/N by A and B, respectively. Then the matrices p^* and q^* with respect to the coordinates $x_1, \cdots, x_r$ are the negative transposes $-A'$ and $-B'$, respectively. The transformation $\overline{p \circ q}$ has then the matrix $AB - BA$ since it is the same as $\bar{q}\bar{p} - \bar{p}\bar{q}$. Hence the transformation of the x's by $(p \circ q)^*$ has the matrix $A'B' - B'A'$, whence we see that this is the same transformation as the transformation by $p^* \circ q^* = q^*p^* - p^*q^*$. Hence $(p^* \circ q^*)\{x_k\} = (p \circ q)^*\{x_k\}$, whence $p^* \circ q^* - (p \circ q)^*$ is singular.

Furthermore, if t is singular and q^* is an arbitrary element of the Lie algebra generated by the p^* with $p \varepsilon P$ then $t \circ q^*$ is also singular. In fact, it is evident that $(t \circ q^*)\{x_i\} = 0$, for $i \leqq r$. Also, if $u \varepsilon L$ we have $q^* \circ u \varepsilon L$, and therefore $u \circ (t \circ q^*) = -t \circ (q^* \circ u) - q^* \circ (u \circ t) = 0$.

Let R denote the Lie algebra which is generated over F by the elements p^* with $p \varepsilon P$. By the above, every element of R is a sum of a p^* and a singular differentiation. Since the set of singular differentiations is of finite dimensions over F, we conclude that R is of finite dimension over F. Let C denote the set of all singular differentiations which belong to R. We have seen that C is an ideal in R, and that $R = P + C$. Moreover, this sum is linearly direct. For, if $p^* + c = p^*_0 + c_0$ then $(p - p_0)^* = c_0 - c$ commutes with every u in L, which implies that, in H, we have $u \circ (p - p_0) = 0$, for every $u \varepsilon L$. But this implies that $p = p_0$, because of the property of P in H. Hence every element $v \varepsilon R$ can be written in the form $v = p^* + c$, where p, not only p^*, is unique. We define the mapping π of R onto P by setting $\pi\{v\} = p$. It is clear that π is a homomorphism of the Lie algebra R onto the Lie algebra P. By Theorem 1.1, there exists an isomorphism α of P into R such that $\pi\alpha$ is the identity mapping on P. Since the kernel of π is C, we have $p^* - \alpha\{p\} \varepsilon C$, whence $u \circ \alpha\{p\} = u \circ p^* = u \circ p$, for every $u \varepsilon L$.

If $h = p + u \varepsilon H$ we define $\beta\{h\} = \alpha\{p\} + u$. Then the mapping β is evidently a homomorphism. Moreover, if $\beta\{h\} = 0$ we have $\alpha\{p\} = -u$,

hence $u_0 \circ p = u_0 \circ (-u)$, for every $u_0 \in L$, which, by the property of P, implies that $p = 0$. Thus, β is an isomorphism.

The differentiations $\beta\{p_i\}$ differ from the t_i only by singular differentiations, and their coefficients are still standard elements. Therefore, all the considerations made in **4** apply also to the $\beta\{p_i\}$. In particular, Theorem 4.1 holds for each $\beta\{p_i\}$. We shall write q_i for $\beta\{p_i\}$.

Now there is a finite dimensional K-linear subspace U of $K\langle x_1, \cdots, x_n\rangle$ which contains the x_i, consists of standard elements, and is mapped into itself by every $u \in L$. By repeatedly applying q_1 to the elements of U we obtain a finite dimensional K-linear subspace $U_1 \supseteq U$ which consists of standard elements and which is mapped into itself by every $u \in L$ and by q_1. Indeed, this is the second assertion of Theorem 4.1. If U_i has already been constructed, we obtain $U_{i+1} \supseteq U_i$ by repeated applications of q_{i+1}. The last of these spaces, U_m, is mapped into itself by every $u \in L$ and by q_m. We claim that, actually, U_m is also mapped into itself by $q_1, \cdots, q_{m-1}$, or, equivalently, that every repeated transform by $q_1, \cdots, q_m$ (not necessarily in order) of an element of U lies in U_m.

We shall prove this by an induction on the total number of the applied differentiations q_i. The result holds trivially if this number is 0 or 1. Suppose it has been established for all q-transforms of elements of U in which the total number of q_i's is less than s. Let v be an s-tuple transform of an element $u \in U$. Since the q_i are images by the isomorphism β of the $p_i \in P$, we have relations $q_j q_i = q_i q_j + \sum_{k=1}^{m} c'_{ijk} q_k$, where the c'_{ijk} are elements of F. By using these relations and the inductive hypothesis it is clear that we can show v to differ from an ordered transform $q^{e_m}{}_m \cdots q^{e_1}{}_1\{u\}$ only by an element in U_m. But the ordered transform belongs to U_m, as is evident from the definition of U_m. Hence $v \in U_m$.

We have proved our main result:

THEOREM 5.1. *Let A be an arbitrary Lie algebra over the field F of characteristic 0. Then A is the direct sum of two Lie algebras T and H, where T is (0) or semisimple, and H can be represented as follows: There is an isomorphism $\bar{h} \to h$ of H onto a Lie algebra of differentiations in $K\langle x_1, \cdots, x_n\rangle$, where n is the dimension of the maximal solvable ideal of H, and K is a finite algebraic extension field of F. There is a finite dimensional subspace W of $K\langle x_1, \cdots, x_n\rangle$ which contains the x_i and which is mapped into itself by every h. If $\tilde{h}$ is the restriction of h to W, the mapping $\bar{h} \to \tilde{h}$ is therefore an isomorphism of H onto a Lie algebra of linear transformations in W, regarding W as a vector space over F.*

Note. Actually, the decomposition $A = T + H$ is no essential restriction. It is easily seen that we can find an isomorphism of A onto a Lie algebra $\bar{A}$ of differentiations in $K\langle x_1, \cdots, x_n; y_1, \cdots, y_s\rangle$, where s is the dimension of T, and there is still a finite dimensional K-linear subspace which yields a faithful linear representation of A. In fact we can arrange matters so that, with W as above, the space $W + y_1W + \cdots + y_sW$ has this property.

UNIVERSITY OF ILLINOIS.

BIBLIOGRAPHY.

[1] I. Ado, "Ueber die Darstellung der endlichen kontinuierlichen Gruppen durch lineare Substitutionen," *Bulletin de la Societé Physico-Mathématique de Kazan*, vol. 7 (1934-1935), pp. 3-43.

[2] G. Birkhoff, "Representability of Lie algebras and Lie groups by matrices," *Annals of Mathematics*, vol. 38 (1937), pp. 526-532.

[3] E. Cartan, "Les representations linéaires des groupes de Lie," *Journal de Mathématiques pures et appliquées*, vol. 17 (1938), pp. 1-12.

[4] C. Chevalley, *Theory of Lie Groups*, Princeton, 1946.

[5] C. Chevalley, "Algebraic Lie algebras," *Annals of Mathematics*, vol. 48 (1947), pp. 91-100.

[6] C. Chevalley and S. Eilenberg, "Cohomology theory of Lie groups and Lie algebras," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 85-124.

[7] Harish-Chandra, "Faithful representations of Lie algebras," *Annals of Mathematics*, vol. 50 (1949).

[8] N. Jacobson, "Restricted Lie algebras of characteristic p," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 15-25.

[9] J. H. C. Whitehead, *Proceedings of the Cambridge Philosophical Society*, vol. 32 (1936), pp. 229-237.

SOME THEOREMS ON ALMOST PERIODIC FUNCTIONS.*

By Raouf Doss.

We shall be concerned here with the following classes of almost periodic functions: the class (B) of Besicovitch, its subclass (Bb) of bounded functions, and the class (B_0) of Bohr functions.

An almost periodic function $f(x) \sim \Sigma b_{u_\nu} e^{iu_\nu x}$ is said to be of basis $\{\beta_i\}$, where the β_i are linearly independent, if each exponent u_ν is a linear combination with rational coefficents of the β_i.

In Theorem I we prove that every linear functional $U(f)$ defined in the space of functions of class (B) and basis $\{\beta_i\}$ is of the form

$$U(f) = \mathfrak{M}\{f(x)\alpha(x)\}$$

where $\alpha(x)$ is a function of the class (Bb).

In Theorem II we prove that every linear functional $U(f)$ defined in the space of functions of class (B_0) and basis $\{\beta_i\}$ is of the form

$$U(f) = \mathfrak{M}\{f(x)\alpha(x)\}$$

where $\alpha(x)$ is a function summable on every finite interval, such that $\mathfrak{M}\{|\alpha(x)|\} < \infty$ and such that $\mathfrak{M}\{e^{iu_\nu x}\alpha(x)\}$ exists for every linear combination u_ν of the β_i.

Theorem I corresponds to the theorem of Steinhaus on the form of the linear functionals defined in the space of summable functions of a given period.

Theorem II corresponds to the case of Riesz. The introduction of the mean value $\mathfrak{M}$ allowed us to dispense with the use of the Stieltjes integral which appears in the Riesz form.[1]

In Theorem III we complete an investigation made some years ago.[2] We give the necessary and sufficient condition that a trigonometric series $\Sigma b_{u_\nu} e^{iu_\nu x}$ should be the development of some function of the class (B). Using an expression of Kovanko, the condition is that the Bochner sums associated with the series are equally B-uniformly summable.

* Received August 2, 1948.

[1] A complete analogue to the Riesz form is given in S. Bochner, "Additive set functions on groups," *Annals of Mathematics*, vol. 40 (1939), pp. 769-799, th. I.

[2] Raouf Doss, "Contribution to the theory of almost periodic functions," *Annals of Mathematics*, vol. 46 (1945), pp. 196-219 (quoted in the sequel as "Contribution").

The method of proof is the same throughout. It is based on the use of the Bochner sums.

1. We fix first our notations. $f(x)$ being summable on every finite interval, we write

$$\bar{\mathfrak{M}}_x\{f(x)\} = \bar{\mathfrak{M}}\{f(x)\} = \limsup\, (2T)^{-1} \int_{-T}^{T} f(x)\,dx$$

as $T \to \infty$, and $\mathfrak{M}\{f(x)\} = \lim\, (2T)^{-1} \int_{-T}^{T} f(x)\,dx$ if this last limit exists.

For any trigonometric series

$$(1) \qquad \Sigma b_{u_n} e^{iu_n x}$$

of basis $\{\beta_i\}$, we define the Bochner sums $\tau_m(x)$ as

$$\tau_m(x) = \sum_{\nu_1=-m!^2}^{\nu_1=m!^2} \cdots \sum_{\nu_m=-m!^2}^{\nu_m=m!^2} (1 - |\nu_1|/m!^2) \cdots (1 - |\nu_m|/m!^2) b_{u_\nu} e^{iu_\nu x}$$

where $u_\nu = \nu_1(\beta_1/m!) + \cdots + \nu_m(\beta_m/m!)$.

We shall write $\tau_m(x) = \Sigma d^m{}_{u_\nu} b_{u_\nu} e^{iu_\nu x}$. If series (1) is the expansion of some function $f(x) \in (B)$ its Bochner sums will be denoted by $f_m(x)$. In that case $f_m(x) = \mathfrak{M}_t\{f(x+t)K_m(t)\}$ where $K_m(t)$ is the Bochner-Fejér Kernel $K_m(t) = \Sigma d^m{}_{u_\nu} e^{-iu_\nu t}$.

LEMMA I. *Let $\sigma(x)$ be a function of Bohr of basis $\{\beta_i\}$ such that l. u. b. $|\sigma(x)| = M$, where $-\infty < x < \infty$. Then the norm of the functional $T(f) = \mathfrak{M}\{f(x)\sigma(x)\}$ defined in the space (B) is $\| T \| = M$.*

Proof. We have evidently $\| T \| \leq M$ since $|T(f)| \leq M \cdot \mathfrak{M}\{|f(x)|\} = M \cdot \| f \|$. On the other hand, let $\epsilon > 0$ be given and let x_0 be such that $|\sigma(x_0)| > (1-\epsilon)M$. The functions $\sigma_m(x) = \mathfrak{M}_t\{\sigma(x+t)K_m(t)\}$ tend uniformly to $\sigma(x)$ as $m \to \infty$. Hence, for sufficiently large m, $|\sigma_m(x_0)| > (1-\epsilon)M$. But

$$\sigma_m(x_0) = \mathfrak{M}_t\{K_m(t-x_0)\sigma(t)\} = \mathfrak{M}_x\{K_m(x-x_0)\sigma(x)\}.$$

We conclude

$$\begin{aligned}(1-\epsilon)M < |\sigma_m(x_0)| &= |T(K_m(x-x_0))| \leq \| T \| \cdot \| K_m(x-x_0) \| \\ &= \| T \| \cdot \mathfrak{M}_x\{K_m(x-x_0)\} = \| T \| \cdot \mathfrak{M}_x\{K_m(x)\} = \| T \| .\end{aligned}$$

ϵ being arbitrary, this gives $\| T \| = M$.

THEOREM I. *Any linear functional $U(f)$ defined in the space of functions of the class (B) and of basis $\{\beta_i\}$ is of the form $U(f) = \mathfrak{M}\{f(x)\alpha(x)\}$ where $\alpha(x) \in (Bb)$.*

Proof. For any linear combination u_n of the β_i put

$$U(e^{iu_n x}) = a_{u_n} = \mathfrak{M}\{e^{iu_n x} a_{u_n} e^{-iu_n x}\}.$$

The Bochner sum associated with the series

$$\Sigma a_{u_n} e^{-iu_n x} \tag{2}$$

will be denoted $\tau_m(x)$. Since $d^m{}_{-u_n} = d^m{}_{u_n}$ we may write

$$\tau_m(x) = \Sigma d^m{}_{u_\nu} a_{-u_\nu} e^{iu_\nu x} = \Sigma d^m{}_{u_\nu} a_{u_\nu} e^{-iu_\nu x}$$

We have, for any $f(x) \sim \Sigma b_{u_\nu} e^{iu_\nu x}$ of the class (B), in virtue of the linearity of U,

$$U(f_m(x)) = \Sigma d^m{}_{u_\nu} b_{u_\nu} a_{u_\nu} = \mathfrak{M}\{f(x)\tau_m(x)\}.$$

If we show that series (2) is the expansion of some function $\alpha(x) \in (Bb)$ with $|\alpha(x)| \leq A$, then

$$U(f_m(x)) = \mathfrak{M}\{f_m(x)\alpha(x)\}.$$

But

$$|\mathfrak{M}\{[f(x) - f_m(x)]\alpha(x)\}| \leq A \cdot \mathfrak{M}\{|f(x) - f_m(x)|\}.$$

Since $\| f - f_m \| = \mathfrak{M}\{|f(x) - f_m(x)|\}$ tend to 0 with $1/m$, then, by the continuity of U

$$U(f) = \lim U(f_m) = \lim \mathfrak{M}\{f_m(x)\alpha(x)\} = \mathfrak{M}\{f(x)\alpha(x)\}, \text{ where } m \to \infty,$$

and our theorem will be proved.

To prove that series (2) is the expansion of some function $\alpha(x) \in (Bb)$, with $|\alpha(x)| \leq A$, it is sufficient to show [3] that

$$|\tau_m(x)| \leq A \tag{3}$$

for every m and x. The linear functionals

$$U_m(f) = U(f_m) = \mathfrak{M}\{f(x)\tau_m(x)\}$$

are weakly convergent, since for every $f \in (B)$ of basis $\{\beta_i\}$, $\lim U_m(f) = U(f)$,

---

[3] See "Contribution," th. VI.

where $m \to \infty$. By a well known theorem of Banach and Steinhaus,[4] their norms $\| U_m \| = \text{u. b.} \mid \tau_m(x) \mid$ are bounded by some constant A. Then (3) is true and the theorem is proved.

2. THEOREM II. *Every linear functional $U(f)$ defined in the space of functions of the class (B_0) and basis $\{\beta_i\}$ is of the form*

$$(1) \qquad U(f) = \mathfrak{M}\{f(x)\alpha(x)\},$$

where $\alpha(x)$ is a function summable on every finite interval, such that $\overline{\mathfrak{M}}\{\mid \alpha(x) \mid\} < \infty$ and such that $\mathfrak{M}\{e^{iu_\nu x}\alpha(x)\}$ exists for every linear combination u_ν of the β_i. Conversely, for any function $\alpha(x)$ with the prescribed conditions, $\mathfrak{M}\{f(x)\alpha(x)\}$ exists and is a linear functional in the space (B_0) of basis $\{\beta_i\}$.

Proof. Put, as before $U(e^{iu_\nu x}) = a_{u_\nu} = \mathfrak{M}\{e^{iu_\nu x} a_{u_\nu} e^{-iu_\nu x}\}$ and let $\tau_m(x)$ be the Bochner sums associated with the series $\Sigma a_{u_\nu} e^{-iu_\nu x}$.

Then, for every $f(x) \in (B_0)$, of basis $\{\beta_i\}$,

$$U(f) = \lim \mathfrak{M}\{f(x)\tau_m(x)\}, \text{ where } m \to \infty,$$

is finite. We conclude[5] that there exists a constant M such that

$$\mathfrak{M}\{\mid \tau_m(x) \mid\} \leq M \quad (\text{for every } m).$$

We know that for every $\tau(x) \in (B_0)$ the integral $(1/L) \int_a^{a+L} \tau(x)\,dx$ tends, uniformly in a, to $\mathfrak{M}\{\tau(x)\}$ as $L \to \infty$. Let $\{\epsilon_n\}$ be a decreasing sequence of positive numbers tending to 0. Since $\mid \tau_m(x) \mid \in (B_0)$, we can determine a sequence $\{L_m\}$ such that for $L \geq L_m$ and every a

$$(2) \qquad 1/L \int_a^{a+L} \mid \tau_m(x) \mid dx < M + \epsilon_m.$$

Similarly, since $e^{iu_\nu x}\tau_m(x) \in (B_0)$, we can determine a sequence $L^\nu{}_1, L^\nu{}_2, \cdots,$ $L^\nu{}_m, \cdots$ such that

$$\left| 1/L \int_a^{a+L} e^{iu_\nu x}\tau_m(x)\,dx - \mathfrak{M}\{e^{iu_\nu x}\tau_m(x)\} \right| < \epsilon_m,$$

i. e.

$$(3) \qquad \left| 1/L \int_a^{a+L} e^{iu_\nu x}\tau_m(x) - d^m{}_{u_\nu} a_{u_\nu} \right| < \epsilon_m.$$

for $L \geq L^\nu{}_m$ and every a.

[4] See, for example, S. Banach, *Théorie des Opérations Linéaires*, Varsovie, 1931, p. 80, th. 5.

[5] See "Contribution," Lemma, p. 210.

Let M_m be the upper bound of $|\tau_m(x)|$. We shall suppose that the following relations, bearing on the diagonal elements $L^m{}_m$ and requiring these elements to be sufficiently large, are satisfied:

$$(4)\quad L_m/L^m{}_m \leqq 1;\quad (5)\quad L^\nu{}_m/L^m{}_m \leqq 1, \nu < m;$$

$$(6)\quad L^\nu{}_m/L^{m-1}{}_{m-1} \leq \epsilon_{m-1},\ \nu < m-1;$$

$$(7)\quad L^\nu{}_m M_m/L^{m-1}{}_{m-1} \leq \epsilon_{m-1},\ \nu < m-1;\quad (8)\quad L_m M_m/L^{m-1}{}_{m-1} \leq \epsilon_{m-1}.$$

Starting from the origin, we shall put, on the right, contiguous intervals $(L^1{}_1), (L^2{}_2), \cdots, (L^n{}_n), \cdots$ of lengths $L^1{}_1, L^2{}_2, \cdots, L^n{}_n, \cdots$ and on the left, contiguous intervals $(-L^1{}_1), (-L^2{}_2), \cdots, (-L^n{}_n), \cdots$ of the same lengths. Let $\alpha(x)$ be a function equal to $\tau_m(x)$ if $x \,\varepsilon\, (L^m{}_m)$ or if $x \,\varepsilon\, (-L^m{}_m)$.

We shall show that $\mathfrak{M}\{|\alpha(x)|\} \leq M + 2\epsilon_1 = M'$.

In fact, let $(-T, T)$ be any interval and suppose that the right end of this interval is in the interval $(L^n{}_n)$, covering a part (l_n) of it, of length l_n. We have

$$\int_0^T |\alpha(x)|\, dx = \int_{(L^1_1)} |\tau_1(x)|\, dx + \cdots + \int_{(L^{n-1}_{n-1})} |\tau_{n-1}(x)|\, dx + \int_{(l_n)} |\tau_n(x)|\, dx.$$

and we have an analogous relation for $\int_{-T}^0 |\alpha(x)|\, dx$.

We consider two cases:

First case. $l_n \geq L_n$. Then by (2) and (4)

$$(9)\qquad \int_0^T |\alpha(x)|\, dx \leq L^1{}_1(M + \epsilon_1) + \cdots + L^{n-1}{}_{n-1}(M + \epsilon_{n-1}) + l_n(M + \epsilon_n) \leq T(M + \epsilon_1).$$

Second case. $l_n < L_n$. Then, by (8)

$$(10)\qquad \int_0^T |\alpha(x)|\, dx \leq [L^1{}_1 + \cdots + L^{n-1}{}_{n-1}](M + \epsilon_1) + l_n M_n \leq T(M + \epsilon_1) + \epsilon_{n-1}L^{n-1}{}_{n-1} \leq T(M + \epsilon_1 + \epsilon_{n-1}).$$

The two relations (9) and (10), together with the analogous relation for $\int_{-T}^0 |\alpha(x)|\, dx$ show that $\mathfrak{M}\{|\alpha(x)|\} \leq M + 2\epsilon_1 = M'$.

We shall now show that, for every u_ν, $\mathfrak{M}\{e^{iu_\nu x}\alpha(x)\}$ exists and is equal to a_{u_ν}.

With the preceding notations, we have, for $n > \nu + 1$, i. e., for sufficiently large T,

$$\int_0^T e^{iu_\nu x}\alpha(x)\,dx = \int_{(L^1_1)} e^{iu_\nu x}\tau_1(x)\,dx + \cdots + \int_{(L^{\nu-1}_{\nu-1})} e^{iu_\nu x}\tau_{\nu-1}(x)\,dx$$
$$+ \int_{(L^\nu_\nu)} e^{iu_\nu x}\tau_\nu(x)\,dx + \cdots + \int_{(L^{n-1}_{n-1})} e^{iu_\nu x}\tau_{n-1}(x)\,dx + \int_{(l_n)} e^{iu_\nu x}\tau_n(x)\,dx.$$

We consider again two cases:

First case. $l_n \geq L^\nu{}_n$. We have, by (3) and (5)

$$(11)\quad \int_0^T e^{iu_\nu x}\alpha(x)\,dx = \int_{(L^1_1)} e^{iu_\nu x}\tau_1(x)\,dx + \cdots + \int_{(L^{\nu-1}_{\nu-1})} e^{iu_\nu x}\tau_{\nu-1}(x)\,dx$$
$$+ L^\nu{}_\nu(d^\nu{}_{u_\nu}a_{u_\nu} + \theta_\nu\epsilon_\nu) + \cdots + L^{n-1}{}_{n-1}(d^{n-1}{}_{u_\nu}a_{u_\nu} + \theta_{n-1}\epsilon_{n-1})$$
$$+ l_n(d^n{}_{u_\nu}a_{u_\nu} + \theta_n\epsilon_n),$$

where $\theta_\nu, \theta_{\nu+1}, \cdots, \theta_n$ have moduli ≤ 1.

Second case. $l_n < L^\nu{}_n$. Then, by (7)

$$\left| \int_{(l_n)} e^{iu_\nu x}\tau_n(x)\,dx \right| \leq l_n M_n < L^\nu{}_n M_n \leq L^{n-1}{}_{n-1}\epsilon_{n-1}.$$

Hence

$$(12)\quad \int_0^T e^{iu_\nu x}\alpha(x)\,dx = \int_{(L^1_1)} e^{iu_\nu x}\tau_1(x)\,dx + \cdots + \int_{(L^{\nu-1}_{\nu-1})} e^{iu_\nu x}\tau_{\nu-1}(x)\,dx$$
$$+ L^\nu{}_\nu(d^\nu{}_{u_\nu}a_{u_\nu} + \theta_\nu\epsilon_\nu) + \cdots + L^{n-1}{}_{n-1}(d^{n-1}{}_{u_\nu}a_{u_\nu} + \theta_{n-1}\epsilon_{n-1}),$$

where $\theta_\nu, \theta_{\nu+1}, \cdots, \theta_{n-2}$ have moduli ≤ 1 and θ_{n-1} a modulus ≤ 2.

Divide each of equations (11) and (12) by T and put

$$\lambda_m{}^T = L^m{}_m/T \ (\nu \leq m < n); \qquad \lambda_m{}^T = l_n/T \ (m = n).$$

As $m \to \infty$, we have $\lim d^m{}_{u_\nu} = 1$, i. e., $\lim (d^m{}_{u_\nu}a_{u_\nu} + \theta_m\epsilon_m) = a_{u_\nu}$, and this limit is uniform in T, since the $\theta_m = \theta_m(T)$ are all in absolute value less than 2.

If we show that the $\lambda_m{}^T$ are the coefficients of a regular process of summation of Toeplitz, then

$$(13)\qquad \lim 1/T \int_0^T e^{iu_\nu x}\alpha(x)\,dx = a_{u_\nu}, \text{ where } T \to \infty.$$

We have to verify two conditions:

1) $\lim \lambda_m{}^T = 0$, where $T \to \infty$ for every m. This is evident.

2) In the first case we have

$$\lim \sum_{m=\nu}^{n} \lambda_m{}^T = \lim \{T - (L^1{}_1 + \cdots + L^{\nu-1}{}_{\nu-1})\}/T = 1, \text{ where } T \to \infty.$$

In the second case we have

$$\sum_{m=\nu}^{n-1} \lambda_m{}^T = \{\mathrm{T} - (L^1{}_1 + \cdots + L^{\nu-1}{}_{\nu-1})\}/T - l_n/T;$$

but, by (6), $l_n/T < L^\nu{}_n/T < L^\nu{}_n/L^{n-1}{}_{n-1} \leqq \epsilon_{n-1}$, so that also

$$\lim \sum_{m=\nu}^{n-1} \lambda_m{}^T = 1, \text{ where } T \to \infty.$$

Relation (13) is thus satisfied; but we have an analogous relation for the integral from $-T$ to 0. Finally

$$\mathfrak{M}\{e^{iu_\nu x}\alpha(x)\} = a_{u_\nu} = U(e^{iu_\nu x}). \tag{14}$$

It remains to be verified that for every $f(x) \in (B_0)$ of basis $\{\beta_i\}$, $\mathfrak{M}\{f(x)\alpha(x)\}$ exists and is equal to $U(f)$.

For an arbitrary $\epsilon > 0$ we can find a large m such that

$$|f(x) - f_m(x)| < \epsilon/(M' + \epsilon), \qquad |U(f) - U(f_m)| < \epsilon. \tag{15}$$

We have, by (14), $U(f_m) = \mathfrak{M}\{f_m(x)\alpha(x)\}$. Take T_0 such that for $T > T_0$, $(2T)^{-1}\int_{-T}^{T} |\alpha(x)|\, dx < M' + \epsilon$ and

$$|U(f_m) - (2T)^{-1}\int_{-T}^{T} f_m(x)\alpha(x)\, dx| < \epsilon. \tag{16}$$

Then

$$|(2T)^{-1}\int_{-T}^{T} [f_m(x) - f(x)]\alpha(x)\, dx| \leqq (\epsilon/M' + \epsilon)\cdot(M' + \epsilon) = \epsilon. \tag{17}$$

The three relations (15), (16), (17) show that, for $T > T_0$,

$$|U(f) - (2T)^{-1}\int_{-T}^{T} f(x)\alpha(x)\, dx| < 3\epsilon.$$

Whence $U(f) = \mathfrak{M}\{f(x)\alpha(x)\}$.

The first part of the theorem is now proved. The second part presents no difficulty.

3. LEMMA 2. *Let $\alpha^n(x) \sim \Sigma a^n{}_{u_\nu} e^{iu_\nu x}$ be a sequence of functions of the*

class (Bb) *of basis* $\{\beta_i\}$ *bounded by the same constant* A *and such that* $\lim a^n{}_{u_\nu} = a_{u_\nu}$, where $n \to \infty$. *Then*

$$\Sigma a_{u_\nu} e^{iu_\nu x} \tag{1}$$

is the expansion of some function $\alpha(x) \epsilon (Bb)$ *and bounded by* A. *Moreover* $\mathfrak{M}\{|\alpha(x)|\} \leq \text{l. u. b.}\, \mathfrak{M}\{|\alpha^n(x)|\} = B$ *and*

$$\mathfrak{M}\{f(x)\alpha(x)\} = \lim \mathfrak{M}\{f(x)\alpha^n(x)\}$$

where $n \to \infty$, *for every* $f(x) \varepsilon (Bb)$ *of basis* $\{\beta_i\}$.

Proof. Let $\epsilon > 0$ be given and let $\alpha_m(x)$ be the Bochner sums associated with (1). We have

$$|\alpha_m(x) - \alpha^n{}_m(x)| \leq \epsilon \qquad \text{(for } m \text{ fixed and } n > n_0). \tag{2}$$

Hence

$$|\alpha_m(x)| \leq |\alpha^n{}_m(x)| + \epsilon \leq A + \epsilon,$$

which proves the first part of the lemma. If now m is chosen such that $\mathfrak{M}\{|\alpha(x) - \alpha_m(x)|\} \leq \epsilon$, then

$$\mathfrak{M}\{|\alpha(x)|\} \leq \mathfrak{M}\{|\alpha^n{}_m(x)|\} + 2\epsilon \leq B + 2\epsilon. \tag{3}$$

Finally, if $|f(x)| < C$ and if m is chosen such that

$$\mathfrak{M}\{|f(x) - f_m(x)|\} < \epsilon, \tag{4}$$

then, for $n > n_0$, by (4) and (2)

$$\begin{aligned} |\mathfrak{M}\{f[\alpha - \alpha^n]\}| &\leq |\mathfrak{M}\{[f - f_m][\alpha - \alpha^n]\}| + |\mathfrak{M}\{f_m[\alpha - \alpha^n]\}| \\ &\leq 2A \cdot \mathfrak{M}\{|f - f_m|\} + |\mathfrak{M}\{f[\alpha_m - \alpha^n{}_m]\}| \leq 2A \cdot \epsilon + C \cdot \epsilon. \end{aligned}$$

This completes the proof.

COROLLARY. *Let* $\alpha^n(x)$ *be functions of the class* (Bb) *of basis* $\{\beta_i\}$, *bounded by the same constant* A. *Then we can extract a partial sequence* $\{\alpha^{n_k}(x)\}$ *and find a function* $\alpha(x) \varepsilon (Bb)$ *bounded by* A, *such that*

$$\mathfrak{M}\{f(x)\alpha(x)\} = \lim \mathfrak{M}\{f(x)\alpha^{n_k}(x)\}, \text{ where } k \to \infty,$$

for every $f(x) \varepsilon (Bb)$ *of basis* $\{\beta_i\}$ *and we have*

$$\mathfrak{M}\{|\alpha(x)|\} \leq \text{l. u. b.}\, \mathfrak{M}\{|\alpha^{n_k}(x)|\}.$$

The partial sequence is obtained by the diagonal process. Note also that by again extracting a subsequence, "upper bound" can be replaced by "limit" and we have

$$\mathfrak{M}\{|\alpha(x)|\} \leq \lim \mathfrak{M}\{|\alpha^{n_k}(x)|\}, \text{ where } k \to \infty.$$

Notation.[6] Given a measurable set E, a summable function $f(x)$ and any interval $(-T, T)$, we mean by $E(-T, T)$ the common part of E and $(-T, T)$ and we write

$$\bar{\delta}E = \lim\sup mesE(-T, T)/2T, \text{ where } T \to \infty,$$

and

$$\bar{\mathfrak{M}}^E{}_x\{f(x)\} = \lim\sup (2T)^{-1} \int_{E(-T,T)} f(x)\,dx, \text{ where } T \to \infty.$$

THEOREM III. *In order that the series $\Sigma a_{u_\nu} e^{iu_\nu x}$ be the expansion of some function $\alpha(x) \in (B)$, it is necessary and sufficient that the Bochner sums $\sigma_m(x)$ attached to the series satisfy the following condition: To every $\epsilon > 0$ there corresponds an $\eta > 0$ such that $\bar{\mathfrak{M}}^E{}_x\{|\sigma_m(x)|\} \leq \epsilon$ for every m and every E for which $\bar{\delta}E \leq \eta$.*

Proof. Necessity. Let the given series be the expansion of $\alpha(x) \in (B)$ and let $\epsilon > 0$ be given. We know that there exists an m_0 such that for every $m \geq m_0$, $\mathfrak{M}\{|\alpha(x) - \sigma_m(x)|\} \leq \epsilon/4$. Whence $\mathfrak{M}\{|\sigma_m(x) - \sigma_{m_0}(x)|\} \leq \epsilon/2$, $(m \geq m_0)$. The polynomial $\sigma_{m_0}(x)$ being bounded, say by the constant A, choose $\eta' = \epsilon/2A$. Hence, for any set E for which $\bar{\delta}E \leq \eta'$,

$$\bar{\mathfrak{M}}^E{}_x\{|\sigma_{m_0}(x)|\} \leq A \cdot \bar{\delta}E \leq A \cdot \epsilon/2A = \epsilon/2,$$

so that, for $m \geq m_0$

$$\bar{\mathfrak{M}}^E{}_x\{|\sigma_m(x)|\} \leq \bar{\mathfrak{M}}_x\{|\sigma_m(x) - \sigma_{m_0}(x)|\} + \bar{\mathfrak{M}}^E{}_x\{|\sigma_{m_0}(x)|\} \leq \epsilon.$$

The number η' therefore verifies the condition of the theorem for $m \geq m_0$. But there is only a finite number of $\sigma_m(x)$ with $m < m_0$ and each of them is bounded. Let A' be an upper bound for $|\sigma_m(x)|$ for $m < m_0$. The number $\eta = \min\{\eta', \epsilon/A'\}$ verifies the condition of the theorem.

Sufficiency. By the condition of the theorem there exists a constant B such that

$$\mathfrak{M}\{|\sigma_m(x)|\} \leq B \qquad \text{(for every } m\text{)}.$$

[6] Kovanko, "Sur la structure des fonctions presque périodiques généralisées," *Recueil Mathématique*, Moscow, vol. 42 (1935), pp. 3-10.

Let n be fixed. Put

$$\sigma^n{}_m(x) = \sigma_m(x) \text{ if } |\sigma_m(x)| \leq n; \quad \sigma^n{}_m(x) = n \cdot \sigma_m(x)/|\sigma_m(x)| \text{ if } |\sigma_m(x)| > n.$$

Then, $\sigma^n{}_m(x) \,\varepsilon\, (B)$.[7]

The functions $\sigma^n{}_m(x)$ form an enumerable set, so that we may suppose that, together with the given series $\Sigma a_{u_\nu} e^{iu_\nu x}$, they belong to the same basis $\{\beta_i\}$. The sequence $\{\sigma^n{}_m(x)\}_m$ being bounded, there exists a partial sequence $\{\sigma^n{}_{m_k}(x)\}_k$ and a bounded function $\sigma^n(x) \,\varepsilon\, (Bb)$ such that

$$\mathfrak{M}\{|\sigma^n(x)|\} \leq \lim \mathfrak{M}\{|\sigma^n{}_{m_k}(x)|\}$$

and

$$\mathfrak{M}\{f(x)\sigma^n(x)\} = \lim \mathfrak{M}\{f(x)\sigma^n{}_{m_k}(x)\}, \text{ where } k \to \infty, \tag{5}$$

for every $f(x) \,\varepsilon\, (Bb)$ of basis $\{\beta_i\}$.

Making use of the diagonal process we might suppose that the same subsequence $\{m_k\}$ is valid for all n and even for all pairs (p, q) since these pairs form an enumerable set, so that we shall write

$$\mathfrak{M}\{|\sigma^p(x) - \sigma^q(x)|\} \leq \lim \mathfrak{M}\{|\sigma^p{}_{m_k}(x) - \sigma^q{}_{m_k}(x)|\}, \text{ where } k \to \infty. \tag{6}$$

Let now E_{m,n_0} be the set of points for which $|\sigma_m(x)| \geq n_0$ and let n_0 be such that $B/n_0 < \eta$. We have $n_0 \cdot \bar{\delta} E_{m,n_0} \leq \bar{\mathfrak{M}}^{E_{m,n_0}}{}_x\{|\sigma_m(x)|\} \leq B$ i. e., $\bar{\delta} E_{m,n_0} \leq B/n_0 < \eta$. Hence

$$\bar{\mathfrak{M}}^{E_{m,n_0}}{}_x\{|\sigma_m(x)|\} \leq \epsilon; \; \bar{\mathfrak{M}}^{E_{m,n_0}}{}_x\{|\sigma^n{}_m(x)|\} \leqq \epsilon. \tag{7}$$

We conclude, by (6), for $p, q > n_0$

$$\begin{aligned}\mathfrak{M}\{|\sigma^p(x) - \sigma^q(x)|\} &\leq \text{l. u. b. } \mathfrak{M}\{|\sigma^p{}_{m_k}(x) - \sigma^q{}_{m_k}(x)|\} \\ &\leq \text{l. u. b. } \bar{\mathfrak{M}}^{E_{m_k,n_0}}{}_x\{|\sigma^p{}_{m_k}(x) - \sigma^q{}_{m_k}(x)|\} \leq 2\epsilon,\end{aligned}$$

since for $x \notin E_{m_k,n_0}$ we have $|\sigma_{m_k}(x)| < n_0$, i. e. $\sigma^p{}_{m_k}(x) = \sigma_{m_k}(x) = \sigma^q{}_{m_k}(x)$. The sequence $\{\sigma^p(x)\}$ is therefore a Cauchy sequence in the space (B). (B) being complete, $\sigma^p(x)$ converges to some function $\alpha(x) \,\varepsilon\, (B)$.

It remains to be shown that the expansion of $\alpha(x)$ is the given series. We can find n_1 such that for $n > n_1$

$$\mathfrak{M}\{|\alpha(x) - \sigma^n(x)|\} < \epsilon. \tag{8}$$

---

[7] See, for example, Kovanko, *loc. cit.*, p. 6.

By (7), for $n > n_0$ and every m_k,

$$(9) \qquad \mathfrak{M}\{|\sigma^n{}_{m_k}(x) - \sigma_{m_k}(x)|\} = \mathfrak{M}^{E_{m_k,n_0}}{}_x\{|\sigma^n{}_{m_k}(x) - \sigma_{m_k}(x)|\} \leq 2\epsilon.$$

By (5), for fixed $n > n_0, n_1$ we can find m_{k_0} such that for $m_k > m_{k_0}$

$$(10) \qquad |\mathfrak{M}\{e^{-iu_\nu x}[\sigma^n(x) - \sigma^n{}_{m_k}(x)]\}| < \epsilon.$$

The three relations (8), (10) and (9) show that for $m_k > m_{k_0}$

$$\begin{aligned} |\mathfrak{M}\{e^{-iu_\nu x}[\alpha(x) - \sigma_{m_k}(x)]\}| &\leq |\mathfrak{M}\{e^{-iu_\nu x}[\alpha(x) - \sigma^n(x)]\}| \\ &+ |\mathfrak{M}\{e^{-iu_\nu x}[\sigma^n(x) - \sigma^n{}_{m_k}(x)]\}| + |\mathfrak{M}\{e^{-iu_\nu x}[\sigma^n{}_{m_k}(x) - \sigma_{m_k}(x)]\}| \\ &\leq \epsilon + \epsilon + 2\epsilon. \end{aligned}$$

Therefore

$$\mathfrak{M}\{e^{-iu_\nu x}\alpha(x)\} = \lim \mathfrak{M}\{\{e^{-iu_\nu x}\sigma_{m_k}(x)\} = \lim d^{m_k}{}_{u_\nu}a_{u_\nu} = a_{u_\nu},$$

where $k \to \infty$.

This completes the proof of the theorem.

Remark 1. We conclude from the preceding theorem that if

$$(11) \qquad f(x) \sim \Sigma b_{u_\nu} e^{iu_\nu x}$$

is an arbitrary function of the class (B) and if $\{\gamma_i\}$ is an arbitrary sequence of linearly independent numbers, then the subseries (Γ) of (11), corresponding to the u_ν which are linear combinations of the γ_i, is the expansion of some function $f^{(\gamma)}(x) \,\varepsilon\, (B)$.[8] For if $H_m(t)$ is the Bochner-Fejér kernel corresponding to $\{\gamma_i\}$ then the sequence $\sigma^{(\gamma)}{}_m(x) = \mathfrak{M}_t\{f(x+t)H_m(t)\}$ is associated to the subseries (Γ) and satisfies the condition of Theorem III. $f^{(\gamma)}(x)$ may be called the restriction of $f(x)$ to the basis $\{\gamma_i\}$.

Remark 2. Theorem I may now be generalized in the sense if $U(f)$ is defined for *all* functions of class (B) then, still $U(f) = \mathfrak{M}\{f(x)\alpha(x)\}$ where $\alpha(x) \,\varepsilon\, (Bb)$.

In fact, put $U(e^{iux}) = a_u$. Then the set of values of u for which $a_u \neq 0$ is at most enumerable. For otherwise, there exists an $a > 0$ such that for a non-enumerable infinity of u: $|a_u| > a$. We can select from these u a sequence of linearly independent numbers $\gamma_1, \gamma_2, \cdots, \gamma_i, \cdots$ such

[8] The corresponding statement for functions of the class (B_0) is due to S. Bochner: "Beiträge zur Theorie der fastperiodische Funktionen," *Mathematische Annalen*, vol. 96 (1926), pp. 119-147.

that $|a_{\gamma_i}| > a$ $(i = 1, 2, \cdots)$. Also $|a_{\gamma_i}| \leq \| U \|$. The function $f(x) \sim \Sigma(1/n)\bar{a}_{\gamma_n}e^{i\gamma_n x}$ belongs to the class (B^2) of Besicovitch and hence to the class (B). If $f_m(x)$ are its Bochner sums we have

$$f_m(x) = \sum_{n=1}^{m} (1 - (1/m!))(1/n)\bar{a}_{\gamma_n}e^{i\gamma_n x}$$

and

$$U(f) = \lim U(f_m) = \lim \sum_{n=1}^{m} (1 - (1/m!))(1/n)|a_{\gamma_n}|^2 = \infty,$$

where $m \to \infty$

which is impossible.

Let then $\beta_1, \beta_2, \cdots, \beta_n, \cdots$ be a basis for the enumerable set of u's for which $a_u \neq 0$. The linear combinations with rational coefficients of the β_i will be denoted $u_1, u_2, \cdots, u_\nu, \cdots$ so that $a_u = 0$ if u is not some u_ν. Then, as in Theorem I $\Sigma a_{u_\nu}e^{-iu_\nu x}$ is the Fourier series of some function $\alpha(x) \varepsilon (Bb)$ and for any $f(x) \varepsilon (B)$ of basis $\{\beta_i\} : U(f) = \mathfrak{M}\{f(x)\alpha(x)\}$. If $f(x)$ has no exponent equal to some u_ν then $U(f) = 0$. Also in that case $\mathfrak{M}\{f(x)\alpha(x)\} = 0$. Finally, if $f(x)$ is any function of class (B), then putting $f(x) = f^{(\beta)}(x) + f(x) - f^{(\beta)}(x)$ we have

$$\begin{aligned} U(f) &= U(f^{(\beta)}) + U(f - f^{(\beta)}) = U(f^{(\beta)}) = \mathfrak{M}\{f^{(\beta)}(x)\alpha(x)\} \\ &= \mathfrak{M}\{f^{(\beta)}(x)\alpha(x)\} + \mathfrak{M}\{[f(x) - f^{(\beta)}(x)]\alpha(x)\} = \mathfrak{M}\{f(x)\alpha(x)\}. \end{aligned}$$

FAROUK I UNIVERSITY,
ALEXANDRIA, EGYPT.

APPLICATION OF A RADICAL OF BROWN AND McCOY TO NON-ASSOCIATIVE RINGS.*

By Malcolm F. Smiley.

1. Introduction. Our first purpose in this paper is to point out that the theory of "radicals" of an associative ring as given by Brown and McCoy [1] applies without change to non-associative rings. We then examine the relation of a particular radical [2] to those defined by A. A. Albert [3] for non-associative algebras and by Max Zorn [4] for hypercomplex alternative rings. We find that the radical [2] of Brown and McCoy is the same as that of Albert for a non-associative algebra *which has a unit element* and also that the radical of Brown and McCoy is the same as Zorn's for a hypercomplex alternative ring.

Our résumé of the theory of Brown and McCoy is preceded by a brief outline of the fundamentals of the theory of non-associative rings and of subdirect sums of such rings. The results of the theory of Brown and McCoy for non-associative rings are then stated without further proof. We conclude our paper with a discussion of the relations between the radical we shall use and those of Albert and of Zorn.

We are indebted to Professor McCoy for directing our attention to his theory and for a stimulating correspondence during the preparation of this paper.

2. Fundamental properties of non-associative rings. In this section we shall briefly outline the facts concerning non-associative rings which we shall need in our ensuing development. All of these facts are well known and we include their statement only for the sake of completeness.

A non-associative ring (*naring*) R is an algebraic system with two single-valued operations $a + b$ and ab defined and in R for every $a, b \in R$ and such that the system $(R, +)$ is an abelian group and the distributive laws $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$ hold for every $a, b, c \in R$. It is easy to prove that $0a = a0 = 0, -(-a) = a, (-a)b = a(-b) = -(ab)$,

* Received September 3, 1948.

[1] Brown and McCoy [1]. See also McCoy [1].

[2] The F_1-radical in the notation of Brown and McCoy [1].

[3] Albert [1].

[4] Max Zorn [2].

and $(-a)(-b) = ab$ for every $a, b \in R$, where 0 and $-a$ denote the unit and the inverse of a, respectively, in the abelian group $(R, +)$.

If R is a naring, then a subset M of R is called an *ideal of* R in case for every $a, b \in M$ and every $x \in R$ we have $a - b$, ax, and $xa \in M$. Then M is empty or M is a subgroup of the additive group $(R, +)$ of R and the cosets $\bar{a} = a + M$ constitute a naring if we define $\bar{a} + \bar{b} = \overline{a + b}$ and $\bar{a}\bar{b} = \overline{ab}$. We call the naring of cosets $\bar{a}$ the *difference naring of* R *and* M and we denote this naring by $R - M$. The mapping $a \rightarrow aH = a + M$ is a homomorphism (the *natural* homomorphism) of R onto $R - M$. If R and $\bar{R}$ are narings and $a \rightarrow aT \in \bar{R}$ is a homomorphism of R onto $\bar{R}$, then the *kernel of* T is the set M of elements $a \in R$ for which $aT = 0$. Then M is an ideal of R and $R - M \cong \bar{R}$ via $a + M \rightarrow aT$. Each ideal N of $R - M$ gives rise to an ideal $N_0 = [x; x + M \in N]$ of R and then $N \cong N_0 - M$. We shall also need the fact that if M and N are ideals of R for which $N \geqq M$, then $(R - M) - (N - M) \cong (R - N)$.

If R is a naring, we shall denote the ideal of R generated by an element $a \in R$ by $I(a)$. More generally, if S is a subset of R we shall denote the ideal of R generated by S by $I(S)$. It is clear that $I(S)$ consists of the set of elements of R of the form $\Sigma s_i + \Sigma s_j U_j$, where the sums are *finite*, s_i, $s_j \in S$, and where each U_j is the product of a finite number of right and left multiplications: $x \rightarrow xR_a = xa$ and $x \rightarrow xL_a = ax$ of R. If $a \rightarrow \bar{a} \in \bar{R}$ is a homomorphism of R onto a naring $\bar{R}$, it is easy to see that $I(\bar{S}) = \overline{I(S)}$.

The ideals of a naring R form a modular lattice when partially ordered by set inclusion. We shall denote the join of two ideals M and N of R by (M, N), while $M \cap N$, as usual, denotes set-theoretic intersection. If M and N are nonvoid ideals of R, then[5] $(M, N) - M \cong N - (M \cap N.)$.

3. Subdirect sums of non-associative rings. For the sake of clarity we shall elaborate in this section the basic theory of subdirect sums as applied to the systems we are considering.

If R_α $(\alpha \in \Omega)$ are narings, then the totality of functions $(a_\alpha; \alpha \in \Omega)$ with $a_\alpha \in R_\alpha$ constitutes a naring S called the *full* direct sum of the narings R_α. A subnaring T of S is called a *subdirect sum* of the narings R_α in case for each $\alpha \in \Omega$ the homomorphism $H_\alpha: a \rightarrow aH_\alpha = a_\alpha$ satisfies $(T)H_\alpha = R_\alpha$.

LEMMA. *A naring* R *is isomorphic to a subdirect sum* T *of narings* $R_\alpha(\alpha \in \Omega)$ *if and only if for each* $\alpha \in \Omega$, R *contains an ideal* M_α *such that* $R - M_\alpha \cong R_\alpha$ *and* $\Pi M_\alpha = 0$.

---

[5] Garrett Birkhoff [1], pp. 47-48.

Proof. Let R be a naring which is isomorphic to a subdirect sum T of narings $R_\alpha(\alpha \varepsilon \Omega)$ via $a \to aH \varepsilon T$. The kernel M_α of HH_α is an ideal of R and $R - M_\alpha \cong R_\alpha$ since $(R)HH_\alpha = (T)H_\alpha = R_\alpha$. To see that $\Pi M_\alpha = 0$, let $a \varepsilon \Pi M_\alpha$, then $aHH_\alpha = 0$ for every $\alpha \varepsilon \Omega$ so that $(aH)_\alpha = 0$, *i.e.*, $aH = 0$, $a = 0$.

Conversely, let $M_\alpha(\alpha \varepsilon \Omega)$ be a set of ideals of R which satisfy the stated requirements. Let h_α denote the natural homomorphism of R onto $R_\alpha \equiv R - M_\alpha$. Then $S \equiv [(\bar{a}_\alpha ; \alpha \varepsilon \Omega) ; \bar{a}_\alpha \varepsilon R_\alpha]$ is the full direct sum of the narings R_α since $(T)H_\alpha = (R)h_\alpha = R_\alpha$. If $a \varepsilon R$, define $aH = (ah_\alpha ; \alpha \varepsilon \Omega) \varepsilon T$. Then $a \to aH$ is an isomorphism of R onto T, since if $aH = bH$, then $(a - b)h_\alpha = 0$ for every $\alpha \varepsilon \Omega$ so that $(a - b) \varepsilon \Pi M_\alpha = 0$ and $a = b$. The proof is complete.

A *subdirectly irreducible* naring R is one which is isomorphic to a subdirect sum T of narings $R_\alpha(\alpha \varepsilon \Omega)$ *only if* H_α is an *isomorphism* for some $\alpha \varepsilon \Omega$. Thus R is subdirectly irreducible if and only if the intersection of all nonzero ideals of R is itself a nonzero ideal J of R. For, if this is true, and if R is isomorphic to a subdirect sum of narings $R_\alpha(\alpha \varepsilon \Omega)$, then, $\Pi M_\alpha = 0$, we have $M\alpha = 0$ for some $\alpha \varepsilon \Omega$ and $R \cong R_\alpha$. On the other hand, if $[M_\alpha ; \alpha \varepsilon \Omega]$ is the totality of nonzero ideals of R, then $\Pi M_\alpha = 0$ would imply that R is isomorphic to a subdirect sum of narings $R_\alpha \cong R - M_\alpha(\alpha \varepsilon \Omega)$. But $R_\alpha \cong R$ is possible for no $\alpha \varepsilon \Omega$ since no M_α is zero. Thus R is not subdirectly irreducible.

4. The F-radical of a non-associative ring. In this and the following section we shall restate the theory of Brown and McCoy as it applies to non-associative rings. We shall omit proofs completely since the proof given by Brown and McCoy are not only valid for non-associative rings but should also be easy for the reader to follow in this case in view of the preparatory material of Sections **2** and **3**.

We assume that $a \to F(a)$ is a mapping defined in each naring[6] R to the set of ideals of R which is such that if $a \to \bar{a} \varepsilon \bar{R}$ is a homomorphism of R onto a naring $\bar{R}$, then $F(\bar{a}) = \overline{F(a)}$. We define the F-radical $N(R, F)$ of R as the set of elements $b \varepsilon R$ such that if $a \varepsilon I(b)$, then $a \varepsilon F(a)$. If $R = N(R, F)$, we call R an *F-radical naring.*

[6] It is possible to phrase our definitions and theorems so as to avoid the metamathematical difficulty of the "class of all rings." We have found this awkward, however, and prefer the present formulation.

THEOREM 1. *The F-radical $N(R, F)$ of a naring R is the intersection of all ideals M of R for which $R-M$ is subdirectly irreducible and $N(R-M, F)=0$.*

COROLLARY 1. *The F-radical of a naring R is an ideal of R.*

COROLLARY 2. *A naring R is an F-radical naring if and only if R itself is the only ideal M of R for which $R-M$ is subdirectly irreducible and $N(R-M, F)=0$.*

THEOREM 2. *If R is a naring, then $N(R-N(R, F), F)=0$.*

THEOREM 3. *If a naring R is a subdirect sum of narings $R_\alpha(\alpha \in \Omega)$ and $N(R_\alpha, F)=0$ for every $\alpha \in \Omega$, then $N(R, F)=0$.*

THEOREM 4. *If R is a naring, then $N(R, F)=0$ if and only if R is isomorphic to a subdirect sum of subdirectly irreducible narings $R_\alpha(\alpha \in \Omega)$ for which $N(R_\alpha, F)=0$.*

THEOREM 5. *A subdirectly irreducible naring R has $N(R, F)=0$ if and only if the minimal ideal J of R contains a nonzero element a such that $F(a)=0$.*

5. The radical of a non-associative ring. In this section we shall discuss the special mapping $a \rightarrow F_1(a)$, where $F_1(a)=I([ax-x+ya-y; x, y \in R])$. Using the results of Section **2**, it is easy to check that this mapping satisfies our basic assumption of Section **4**. We shall write N for $N(R, F_1)$ and we shall call N the radical of the naring R. If $N=0$, we shall say that the naring R is *semi-simple* and if $N=R$ we shall call the naring R a *radical naring.*

THEOREM 6. *A subdirectly irreducible naring is semi-simple if and only if it is a simple naring with unit element.*

Proof. Excluding the trivial case of a one-element ring, the direct statement follows from Theorem 5 since $F_1(1)=0$. Conversely, by Theorem 5, there is a nonzero element e of the minimal ideal J of R for which $F_1(e)=0$. Then e is the unit element of R and $J=R$ so that R is simple.

COROLLARY. *A simple naring is semi-simple if it has a unit element, otherwise it is a radical naring.*

Remark. In the associative case, one is able to show that a simple ring with left (right) unit has a unit, and consequently that the G-, F_1-, and

$F_1{}^r$-radicals all coincide. This result is no longer true in the non-associative case, as the following example shows. Let A be the non-associative algebra of order two over a field F with basal units e and u and multiplication defined by $ee = e$, $eu = u$, $ue = 0$, and $uu = e$. This algebra is simple since a nonzero ideal M of A contains a nonzero element $\alpha e + \beta u$ and hence also contains $u(\alpha e + \beta u) = \beta e$. Thus M contains e unless $\beta = 0$, but then α is nonzero and M contains αe, so that $e \varepsilon M$ in any case. Thus $M = A$ and A is a simple non-associative algebra with left unit but no unit. The existence of a unit for simple narings with one sided units may be obtained without the full force of the associative law. We may, for example, assume that association is *symmetric* in the sense that if $a(bc) = (ab)c$, then a, b, c associate in any order. Then if R is a simple naring with a left unit and if association is symmetric in R, R has a unit element. An example of a naring in which association is symmetric is any *alternative* ring.

THEOREM 7. *The radical of a naring R is the intersection of all the ideals M of R such that $R - M$ is a simple naring with unit element.*

COROLLARY 1. *If R is not a radical naring then the radical of R is the intersection of all the maximal ideals M of R such that $R - M$ has a unit element.*

COROLLARY 2. *If R is a naring which is not a one-element ring and has a unit element, then the radical of R is the intersection of all the maximal ideals of R.*

THEOREM 8. *The naring R is semi-simple if and only if it is isomorphic to a subdirect sum of simple narings with unit element.*

THEOREM 9. *If the descending chain condition holds for the ideals of a semi-simple naring R, then R is isomorphic to the full direct sum of a finite number of simple narings with unit elements.*

Remark. We postpone a discussion of Theorem 10 of Brown and McCoy with the remark that the Jacobson radical has been defined only for associative and alternative rings.[7]

THEOREM 11. *If R is a power-associative*[8] *naring, that is, if each element of R generates an associative subnaring of R, and if every element of $I(b)$ is nilpotent, then b is in the radical of R.*

---

[7] N. Jacobson [1] and Smiley [1].

[8] A. A. Albert [2].

THEOREM 12. *If A is an ideal of the naring R, the radical of A is contained in the radical of R.*

COROLLARY. *Any ideal of a semi-simple naring is semi-simple.*

6. Relation to the radical of A. A. Albert. In this and the following section we shall discuss the relation of the radical of a non-associative ring which was defined in Section **6** to previous definitions of the radical of certain non-associative systems given by A. A. Albert[3] and by Max Zorn.[4]

A. A. Albert calls a non-associative algebra A of finite order over a field "semi-simple" in case A is the direct sum of (finitely many) nonzero simple non-associative algebras. Then if A is homomorphic to a "semi-simple" non-associative algebra, Albert definies the radical N' of A to be the intersection of the family of ideals B_α of A for which $A - B_\alpha$ is "semi-simple." We shall show in this section that $N' = N$ when the basic ring R is a non-associative algebra of finite order *with a unit element.* We emphasize that all of our theorems are valid for non-associative algebras.

LEMMA. *If B is an ideal of a non-associative algebra A such that $A - B$ is isomorphic to the direct sum of nonzero simple non-associative algebras C_i $(i = 1, \cdots, n)$, then there are ideals B_i of A $(i = 1, \cdots, n)$ such that each $A - B_i$ is a nonzero simple non-associative algebra and $\Pi B_i = B$.*

Proof. The mapping $a \rightarrow aH_i = c_i$ is a homomorphism of $A - B$ onto C_i with kernel M_i. Now M_i is an ideal of $A - B$ and $(A - B) - M_i \cong C_i$. Each M_i gives rise to an ideal B_i of A such that $B_i \geqq B$ and $M_i = B_i - B$. Hence $A - B_i \cong (A - B) - (B_i - B) = (A - B) - M_i \cong C_i$, so that $A - B_i$ is a nonzero simple non-associative algebra. To see that $\Pi B_i = B$, note that $\Pi M_i = 0$ from which $\Pi B_i = B$ follows readily.

COROLLARY 1. *If A is a non-associative algebra which is homomorphic to a "semi-simple" non-associative algebra, then N' is the intersection of all the ideals C_β of A for which $A - C_\beta$ is a nonzero simple non-associative algebra.*

COROLLARY 2. *If A is a non-associative algebra with unit element then N' is the intersection of all ideals C_α of A for which $A - C_\alpha$ is a simple non-associative algebra with unit element.*

Proof. It suffices to remark that if $A - C_\alpha$ is a one element algebra, then $A = C_\alpha$.

THEOREM 13. *If A is a non-associative algebra of finite order over a field and which has a unit element, then $N' = N$.*

Proof. This is clear by Corollary 2 of the Lemma and Theorem 7.

Albert has given an example [8] for which N' is a *field* provided that the characteristic of the base field is *not two.* In this exceptional case the associative subalgebra spanned by e and u contains the nilpotent ideal generated by $e + u$ and *is not* semi-simple. We find that N' is the ideal of A spanned by $e + u$ and v.

7. The radical of a hypercomplex alternative ring. We continue our discussion of previous work on the radical of certain non-associative systems. Max Zorn [9] proved that if an alternative ring R was a *hypercomplex* alternative ring, then the set of properly nilpotent elements of R is an ideal and further if this ideal is zero for R, then R is the direct sum of a finite number of simple alternative rings (with unit). We have shown elsewhere [10] that Jacobson's definition of the radical of a ring [9] applies to alternative rings and is the set of all properly nilpotent elements provided that the alternative ring is a hypercomplex alternative ring. In this section we shall show that the radical of Brown and McCoy also reduces to the set of all nilpotent elements for a hypercomplex alternative ring.

A *hypercomplex* alternative ring is an alternative ring which satisfies the following chain conditions.

(CI) *Every sequence* $(a^nR; n = 1, 2, \cdots)$ *is ultimately constant.*

(CII) *Every monotone sequence* $(A_r(a_n); n = 1, 2, \cdots)$ *is ultimately constant.*

Here $A_r(a_n) \equiv [x; x \varepsilon R, a_nx = 0]$ is the set of all right annihilators of a_n. We then have the following theorem.

THEORM 10. *If R is a hypercomplex alternative ring, then N is the set of all properly nilpotent elements of R, that is, N is the radical of R in the sense of Zorn.*

Proof. We have shown elsewhere [10] that under the hypothesis of our theorem, the set of all properly nilpotent elements of R coincides with the radical of R as defined by Jacobson. This latter radical consists of all

[9] Max Zorn [1] and [2].
[10] Smiley [1].

elements $b \varepsilon R$ such that $a \varepsilon I(b)$ implies that there is an element $c \varepsilon R$ so that $a - c + ac = 0$. We have also shown that $a - c + ac = 0$ for a, $c \varepsilon R$ if and only if every element of R is in the set $[ax - x; x \varepsilon R]$. Then clearly N contains the radical N_2 of Zorn.

To prove that the radical of Zorn contains N we use the result of Zorn which states that N_2 is an ideal of R. Observe that $R - N_2$ is a hypercomplex alternative ring and that $N_2(R - N_2) = 0$. Then $R - N_2$ is a direct sum of simple alternative rings with unit elements. Thus Theorem 8 yields $N(R - N_2) = 0$, from which $N_2 \geqq N$ follows easily. The proof is complete.

NORTHWESTERN UNIVERSITY.

BIBLIOGRAPHY.

A. A. Albert:

[1] "The radical of a non-associative algebra," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 891-897.

[2] "Power associative rings. I," (Abstract), *Bulletin of the American Mathematical Society*, vol. 53 (1947), p. 905.

Garrett Birkhoff:

[1] *Lattice Theory*, New York City, 1940.

B. Brown and N. H. McCoy:

[1] "Radicals and Subdirect Sums," *American Journal of Mathematics*, vol. 69 (1947), pp. 46-58.

N. Jacobson:

[1] "The radical and semi-simplicity for arbitrary rings," *American Journal of Mathematics*, vol. 67 (1945), pp. 300-320.

N. H. McCoy:

[1] "Subdirect sums of rings," *Bulletin of the American Mathematical Society*, vol. 53 (1947), pp. 856-877.

M. F. Smiley:

[1] "The radical of an alternative ring," *Annals of Mathematics*, vol. 49 (1948), pp. 702-709.

M. Zorn:

[1] "Theorie der alternativen Ringe," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 8 (1930), pp. 123-147.

[2] "Alternative rings and related questions I: Existence of the radical," *Annals of Mathematics* (2), vol. 42 (1941), pp. 676-686.

ON n-ALITY THEORIES IN RINGS AND THEIR LOGICAL ALGEBRAS, INCLUDING TRI-ALITY PRINCIPLE IN THREE VALUED LOGICS.* [1]

By ALFRED L. FOSTER.

Introduction. In every ring $(R, +, \times)$ there exists an intrinsic but usually dormant duality-symmetry theory which specializes to the familiar Boolean duality when R is a Boolean ring. This theory has been presented and developed in diverse directions in a series of papers [1], $\cdots$, [7],[2] with several of which it will be necessary to establish contact in the present communication.

It was later discovered, as first broadly outlined in a portion of [2], that this duality theory of rings is itself but an instance of a host of K-ality theories, based on certain preassigned groups K, and that these in turn constitute merely one class of realizations of a general transformation theory,—a simple unifying skeletal framework which also includes traditional transformation and invariant theories among its specializations.

Even on the *simple* (later also called mod C) duality level, as we now refer to the original ring duality to distinguish it from rival theories, we were able to formulate numerous interesting concepts, such as the (simple) 'logical algebra' of a ring, and to profitably explore such questions as the strength of the bond between a ring and its (simple) logic; these include generalizations of familiar Boolean questions. (See especially [1]). In the present communication we (a) elaborate the general transformation theory and employ it to (b) elevate various 'logical algebra' concepts from the simple to a much more general level. In so doing (c) the special role of the simple level is considerably illuminated.

In **6** we (d) put forward the concept 'p-ring' as a natural generalization of Boolean rings (which latter are identical with 2-rings),[3] and particularly

* Received November 1, 1948.

[1] A segment of this paper was presented to the National Academy of Sciences, Berkeley, Nov. 1948.

[2] Numbers in square brackets refer to the bibliography given at the conclusion.

[3] The concept "p-ring" was first defined by McCoy and Montgomery, in [9]—for which reference I am indebted to the referee. Strictly, our "p-ring" is the "p-ring with unit" of [9].

study the case $p=3$. It is shown that each 3-ring is interdefinably bound to its 'logical algebra' in a manner which generalizes the familiar interdefinable bond between a Boolean ring (i. e., $p=2$) and its Boolean (= logical) algebra. Each 3-ring-algebra is shown to possess an intrinsic *tri-ality* theory, the successor of the familiar Boolean duality case for $p=2$. In this way the same tight intimacy which exists between the logic of propositions (= 2-valued logic), Boolean rings, Boolean algebras and the omnipresent Boolean duality theory on the one hand, is shown to extend to the 3-valued logic, 3-rings, the corresponding 3-algebras, and the engulfing tri-ality theory on the other.

For arbitrary p (= prime), we exhibit a *p-ality* theory connecting each *p-ring* with its logical alegbra. For $p>3$, however, no formula has as yet been found which equationally defines a p-ring in terms of its logical algebra, in fact the existence of such a formula has not been settled,—(see **8-10**).

1. (Simple) Duality theory of rings. To facilitate reference and orientation, in this section we very briefly recall a few essentials of the simple duality theory of rings.

Let $R=(R,+,\times)$ be a ring with a unit element. Each concept in R is shown to possess a dual concept; in particular: 0 and 1 are dual elements; $\times,\otimes$; $+,\oplus$; $-,\ominus$; * are dual operations, the latter being self-dual, where

(1.1) $\quad a\otimes b=a+b-a\times b,\qquad a\times b=a\oplus b\ominus a\otimes b$ = dual ring products

(1.2) $\quad a\oplus b=a+b-1,\qquad a+b=a\oplus b\ominus 0$ = dual ring sums

(1.3) $\quad a\ominus b=a-b+1,\qquad a-b=a\ominus b\oplus 0$ = dual ring differences

(1.4) $\quad a^*=1-a=0\ominus a=$ (self-dual) ring complement.

Restricted for brevity to these operations the duality theorem reduces to:

(Simple) Duality Theorem for Rings. *If* $P(0,1;\times,\otimes;+,\oplus;-,\ominus;{}^*)$ *is a true proposition of a given ring R, so also is its (simple) dual,* $\mathrm{dl}\,P=P(1,0;\otimes,\times;\oplus,+;\ominus,-;{}^*)$ *obtained by replacing each argument by its dual, with * left unchanged (self dual).*

The duality theorem is illustrated by each of the relations (1.1)-(1.4). Also by the dual theorems, holding in any ring,

$$(1.5) \qquad (a \times b)^* = a^* \otimes b^*, \qquad (a \otimes b)^* = a^* \times b^*$$
$$= \text{Ring ‘De-Morgan’ formulas,}$$

and the self dual

$$(1.6) \qquad a^{**} = a, \qquad 0^* = 1, \qquad 1^* = 0.$$

Again, by

$$(1.7) \qquad \begin{matrix} (R, +, \times) \text{ is a ring with 0 as zero element and 1 as unit.} \\ (R, \oplus, \otimes) \text{ is a ring with 1 as zero element and 0 as unit.} \end{matrix}$$

Also, in any Boolean-like ring, (see [1]),

$$(1.8) \qquad a + b = (a \times b^*) \otimes (a^* \times b), \qquad a \oplus b = (a \otimes b^*) \times (a^* \otimes b).$$

Again, in any field $(F, +, \times)$, (see [6]),

$$(1.9) \qquad \begin{matrix} a \times a^{01} \times a^{10} = \mu = \text{constant} \neq 0 & (a \neq 0, \neq 1) \\ a \otimes a^{10} \times a^{01} = \tau = \text{constant} \neq 1 & (a \neq 1, \neq 0) \end{matrix}$$

$$(1.10) \qquad a + b : \begin{cases} a + b = a \otimes b \otimes (a^0 \times b^0) & (a \neq 1, b \neq 1) \\ a + 1 = 1 + a = a \otimes a^{10} & (a \neq 0, \neq 1) \\ 1 + 1 = \tau, \qquad 1 + 0 = 0 + 1 = 1 & \end{cases}$$

$$(1.11) \qquad a \oplus b : \begin{cases} a \oplus b = a \times b \times (a^1 \otimes b^1) & (a \neq 0, b \neq 0) \\ a \oplus 0 = 0 \oplus a = a \times a^{01} & (a \neq 1, \neq 0). \\ 0 \oplus 0 = \mu, \qquad 0 \oplus 1 = 1 \oplus 0 = 0 & \end{cases}$$

In (1.9) and (1.10), a^1 *is the* $\times$ inverse of a, and a^0 the $\otimes$ inverse. These few illustrations will suffice for our purpose. In case R is a Boolean ring we recall that: the dual ring products $\times$, $\otimes$ respectively reduce to the usual Boolean logical product, $\cap$, and logical sum, $\cup$, the ring complement $*$ to the Boolean complement, $^-$, and the ring ‘De Morgan’ and duality theorems to the corresponding familiar Boolean theorems.

With the above Boolean specialization as motivation, in any ring R the (operationally closed) system $(R, \times, \otimes, *)$ was introduced in [1] as the (simple) *logical algebra* (also briefly as the *Logic*) of the ring. We recall that a *Logically definable* ring is one whose ring $+$, and consequently the entire ring, is definable in terms of its Logic, as is the case for instance in a Boolean ring. These notions will be reexamined and refined in **5**.

In addition to the usual $(R, +, \times)$ notation for a ring we may also write it in the (simple) dual form $(R, \oplus, \otimes)$, or in the ‘mixed’ form $(R, +, \oplus, \times, \otimes, *)$, etc. Later, corresponding to other duality or, more generally, n-ality theories, we have n pure forms

$$(R, +, \times), \qquad (R, +', \times'), \qquad (R, +'', \times''), \cdots$$

and corresponding 'mixed' forms

$$(R, +, +', +'', \cdots, \times, \times', \times'', \cdots, {}^{\circ}, {}^{\circ\prime}, {}^{\circ\prime\prime}, \cdots).$$

2. Perspective of the general transformation theory. In the traditional applications of the usual transformation and invariant theories to various mathematical disciplines, basically one is concerned with a set (generally a group) of admissible 'coordinate transformations,' and with the changes suffered by (or with the invariance of) certain mathematical concepts of the discipline when one passes from one admissible coordinate system to another such. In all of the classical applications the underlying 'computational disciplines' (arithmetic, various algebras, analysis etc.) are absolute invariants (scalars), i. e., unchanged by any of the coordinate transformations. In [2] it was first sketched how this transformation-invariant theory may be profitably extended to permit much wider applications, as for instance to the above types of computational disciplines themselves. Thus one is led to the conception of arithmetic, or analysis, or a particular ring or class of rings, in fact any kind of operational algebra, as a discipline whose concepts transform 'cogrediently,'—or sometimes 'contragrediently' with each permissible change of 'coordinates,' in a manner analogous to (and possessing as a particular specialization) the transformation of tensors or matrices to new coordinates.

In general one is forced to deal with 'mixed' as well as 'pure' notions, the latter being such as are defined entirely within a single coordinate system, while the former are defined in terms of at least two permissible systems. Thus, for example, while a ring defined in either of the (simple) dual forms $(R, +, \times)$ or $(R, \oplus, \otimes)$ is a pure concept, the Logic $(R, \times, \otimes, {}^*)$ of the ring is perforce a 'mixed' notion.

For each given group K of permissible 'coordinate transformations' in a discipline there is a *K-ality theorem* which explicitly formulates the manner in which the true propositions of the discipline transform when passing from one permissible coordinate system to another. Thus for instance, corresponding to each group K of order 2 one has an accompanying duality theory; in particular, applied to a ring R and with K chosen as the *complementation group* $C = C(R)$, of order 2,

$$x^* = 1 - x, \qquad x^{**} = x = \text{identity}, \tag{2.1}$$

the corresponding ring duality is the *simple* theory partially recalled in 1.

From the general transformation theory point of view, then, the 'K-ality' of concepts,—also variously referred to as 'n-ality, mod K,' or 'n-ality (K),' etc., if K is a finite group of order n, is simply a way of saying that the concepts are identical, but expressed in different coordinates. Thus for example, in the (simple) ring duality theory specialized to Boolean rings, the familiar logical product $\times\ (=\cap)$ and logical sum $\otimes\ (=\cup)$ are exhibited as the *same* concept, expressed in different coordinates.

With the exception of [4] and parts of [2], the narrow band of applications of this transformation theory explored in the series [1], $\cdots$, [7] is more or less identified with rings, and with the permissible group K taken as the simple complementation group C. We shall here still be concerned with rings, but not alone with this special choice of the group K.

3. General transformation theory (continued). Let $U=\{\cdots,x,\cdots\}$ be a class (with or without structure), and $\Phi_U=\Phi=\{\cdots,\phi,\cdots\}$ the set of all operations (or *multitations*,—see [4])

$$\phi=\phi(x,\cdots)\ \varepsilon\ U \tag{3.1}$$

(or one or more arguments) of U into itself. A ϕ of a single argument is also called a *monotation*, $\phi(x)$; similarly for *bitations* $\phi(x,y)$, etc. We have

$$\Phi=\sum_{i=1,2,\dots}\Phi^{(i)} \tag{3.2}$$

where $\Phi^{(i)}$ is the class of all i-tations of U, and $\sum$ denotes set union. A permutation ρ is a 1-1 reversible monotation,[4] whose inverse we write $\bar{\rho}$. For $\sigma\ \varepsilon\ \Phi^{(i)}$ and $\phi\ \varepsilon\ \Phi$ we denote composition ($=$composite product) by simple juxtaposition,

$$\sigma\phi=(\sigma\phi)(x,\cdots)=\sigma(\phi(x,\cdots)). \tag{3.3}$$

We recall the well known associativity of composition for monotations,

$$\sigma(\sigma'\sigma'')=(\sigma\sigma')\sigma''=\sigma\sigma'\sigma''. \tag{3.4}$$

Solely in the interest of convenient distinction we sometimes refer to the elements of the set U as 'points,' and those of the set Φ_U as Points (capital P): $U=$'point space,' $\Phi=$'Point space.'

---

[4] In particular, U need not, of course, be infinite.

Each point permutation ρ induces a Point monotation, defined by

$$\phi \to \phi_\rho = \rho^{-}\phi(\rho(x), \rho(y), \cdot \cdot \cdot). \tag{3.5}$$

The Point ϕ_ρ we call the *transform of* ϕ *by* ρ.

Thus, if U is chosen as a linear vector space $U = \{\cdot \cdot \cdot, \bar{x}, \cdot \cdot \cdot\}$, ϕ as a linear map

$$\phi(\bar{x}) = \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots \\ \phi_{21} & \phi_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

and ρ as a non-singular linear map, the transform (3.5) reduces to the familiar matrix transform $\rho^{-}\phi\rho$. Similar familiar interpretations result if U is taken as an abstract group. Again if U is taken as a ring $(U, +, \times)$ with unit, ϕ as the ring product $\times$,

$$\phi(x, y) = x \times y \tag{3.6}$$

and ρ as the *complementation* permutation

$$\rho(x) = x^* = 1 - x = \rho^{-}(x), \tag{3.7}$$

the ρ transform of $\times$ is the simple dual, $\otimes$, given by (1.1); and quite generally, the ρ transform of any concept of U is its simple dual. (Compare with 1).

THEOREM 1. (a) *Each induced Point monotation* (3.5) *is a Point permutation.* (b) *If* $K = \{\cdot \cdot \cdot, \rho, \cdot \cdot \cdot\}$ *is a group of point permutations, the corresponding set of induced Point permutations* (3.5) *form a group, which is isomorphic with the group* K. *One has*

$$\{\phi\}_{\rho\rho'} = \{\phi_\rho\}_{\rho'}. \tag{3.8}$$

Proof of (a). Here we must show that (a'): (3.5) is a univoque Point monotation, i. e.,

$$\phi \neq \psi \to \phi_\rho \neq \psi_\rho, \tag{a'}$$

and secondly (a''): for each Point ψ there exists a Point ψ' such that $\psi'_\rho = \psi$. If (a') were false there would exist points $x_0, y_0, \cdot \cdot \cdot$ such that

$$\phi(x_0, y_0, \cdot \cdot \cdot) \neq \psi(x_0, y_0, \cdot \cdot \cdot) \tag{3.9}$$

$$\rho^{-}\phi(\rho(x), \rho(y), \cdot \cdot \cdot) \underset{x, y, \ldots}{\equiv} \rho^{-}\psi(\rho(x), \rho(y), \cdot \cdot \cdot). \tag{3.10}$$

This is impossible since (3.10) implies

$$\phi(\rho(x),\rho(y),\cdot\cdot\cdot)\underset{x,y,\ldots}{\equiv}\psi(\rho(x),\rho(y),\cdot\cdot\cdot), \tag{3.11}$$

which is in contradiction with (3.9), as is seen by taking $x=x_1$, $y=y_1,\cdot\cdot\cdot$ where

$$\rho(x_1)=x_0, \qquad \rho(y_1)=y_0,\cdot\cdot\cdot. \tag{3.12}$$

This proves (a′). Again, for given $\psi\,\varepsilon\,\Phi$, (a″) may be satisfied by defining ψ' by

$$\psi'(x,y,\cdot\cdot\cdot)=\rho(\psi(\rho^{-}(x),\rho^{-}(y),\cdot\cdot\cdot)). \tag{3.13}$$

This completes part (a).

The relation (3.8) follows directly from the usual expression for the inverse of a product:

$$\begin{aligned}\{\phi_\rho\}_{\rho'}(x,y,\cdot\cdot\cdot)&=\rho'^{-}(\rho^{-}(\phi(\rho(\rho'(x)),\rho(\rho'(y)),\cdot\cdot\cdot)))\\ &=(\rho\rho')^{-}(\phi(\rho\rho'(x),\rho\rho'(y),\cdot\cdot\cdot))=\phi_{\rho\rho'}.\end{aligned} \tag{3.14}$$

Part (a) and (3.8) together show that the set of induced Point monotations form a group which is a homomorphic image of the group K. That we actually have an isomorphic image may be seen as follows. One must show that for $\rho\neq\rho'$, (b′): $\phi\rightarrow\phi_\rho$ and $\phi\rightarrow\phi_{\rho'}$ are different Point permutations. From the premise, for some $x_0\,\varepsilon\,U$,

$$\rho(x_0)\neq\rho'(x_0). \tag{3.15}$$

Denote

$$\rho(x_0)=x_1, \qquad \rho'(x_0)=x_2. \tag{3.16}$$

Then $x_1\neq x_2$. Consider the Point $\psi(x)\underset{x}{\equiv}x_1$. One has

$$\begin{aligned}\psi&\rightarrow\psi_\rho=\rho^{-}\psi(\rho(x))\underset{x}{\equiv}\rho^{-}(x_1)=x_0\\ \psi&\rightarrow\psi_{\rho'}=\rho'^{-}\psi(p'(x))\underset{x}{\equiv}\rho'^{-}(x_1).\end{aligned} \tag{3.17}$$

Now $\rho'^{-}(x_1)\neq x_0$, since

$$\rho'^{-}(x_1)=x_0\rightarrow x_1=\rho'(x_0)=x_2, \tag{3.18}$$

contrary to hypothesis. Hence (b′) is proved, and with it the Theorem.

The group of Point transformations induced by a group K will not in general be transitive, even if the point transformation group is transitive.

The totality of Points ϕ_ρ into which a given Point ϕ is transformed by the various $\rho \varepsilon K$ is called a *congruence class* (of Points), mod K. We write

$$\phi \equiv \phi'(K) \tag{3.19}$$

to denote that ϕ and ϕ' are in the same congruence class. Since each congruence class mod K forms a Point set on which the group K is homomorphically represented as a transitive permutation group, by a well known theorem on the degree of a transitive permutation group one has the

THEOREM 2. *If the group K is finite, of order n, the number of Points in a congruence class, mod K, is a divisor of n.*

In particular, if K is a (cyclic) group of prime order p, the number of Points in each congruence class mod K is either p or 1.

Since $K \subset \Phi$, each $\rho \varepsilon K$ is of course a Point. One has

THEOREM 3. *If K is an Abelian group, each $\rho \varepsilon K$ is fixed under K, i. e., forms a congruence class of a single Point.*

The proof is immediate from (3.5). The self-duality of the ring complement, *, in the simple ring duality theory, and similarly the self-tri-ality of the cyclical negation, $^\wedge$, as well as its inverse, $^\vee$, in the tri-ality theory of 3-rings considered in **6-10**, is an immediate consequence of Theorem 3. Of course there may also be fixed Points ϕ, mod K, which are not elements of K.

It is often advantageous to regard these general transformation notions in a different light, corresponding to **2**. We may look on K as a group of permissible 'coordinate transformations' in U. In the 'ρ coordinate system' the point x receives the new 'coordinate' $\rho(x)$ [5]; the multiplication ϕ 'becomes' ϕ_ρ. That is, ϕ_ρ, which is an isomorphic image of ϕ by Theorem 1, is the 'same' multitation as ϕ, described however in the ρ coordinate system.

For a given group K of permissible coordinate transformations we also say: ϕ and ϕ' are 'K-als,' instead of $\phi \equiv \phi'(K)$. Moreover if K is finite, of order n, we also speak of 'n-als' mod K; for $n = 2, 3, \cdots$ we speak of 'duals,' 'tri-als,' etc. A fixed Point ϕ is then called 'self-K-al,' respectively 'self-dual,' 'self-tri-al,' etc.

If $A = \{\alpha_1, \alpha_2, \cdots\}$ is any subset of Φ_U, we denote by $|A|$ the subset of Φ_U which is compositionally generated by the $\alpha_i \varepsilon A$. For instance if

[5] More accurately, by the 'point x' we mean the point which, in the ξ coordinate system, has the coordinate x, where ξ is the identity of the group K.

$A = \{\alpha, \beta\}$, and if α, β are of the form $\alpha(x), \beta(x, y)$, then $|A|$ would contain the multitations

$$\alpha(\alpha(x)), \alpha(\beta(x, y)), \beta(\alpha(x), \beta(y, z)), \beta(\alpha(x), \beta(x, y)), \text{ etc.} \tag{3.20}$$

We have of course

$$A \subseteq |A| \subseteq \Phi\Phi_U. \tag{3.21}$$

4. *K*-logical definability of rings. We here clarify the concept of the simple logical (algebra) definability of rings (see end of **1**), and lift it from the simple to the general level. We specialize: $U = R = (R, +, \times)$ is a ring (which need not contain a unit). Let

$$K = \{\xi, \rho', \rho'', \cdots\} = \{\cdots, \rho, \cdots\} \tag{4.1}$$

be a group of coordinate transformations in (i. e., permutations of) the class R, with the identity of K denoted by ξ, and let

$$\{\times, \times', \times'', \cdots\} = \{\cdots, \times_\rho, \cdots\} \tag{4.2}$$

be the class of all transforms of the ring product $\times$ by the various $\rho \varepsilon K$. (This congruence class mod K is evidently the same as the class of transforms of any fixed $\times^{(4)}$). Here, by (3.5),

$$a \times_\rho b = \rho^{-}(\rho(a) \times \rho(b)), \tag{4.3}$$

and obviously

$$a \times_\xi b = a \times b. \tag{4.4}$$

The algebra

$$(R, \times, \times', \times'', \cdots, \xi, \rho', \rho'', \cdots) \tag{4.5}$$

whose class, R, is taken as identical with that of the given ring $(R, +, \times)$, and whose operations are $\times, \times', \times'', \cdots, \xi, \rho', \rho'', \cdots$ as indicated, we call the *logical alegbra* (mod K of the ring, or simply the *K-logic* of the ring. In addition to (4.5) we also denote this K-logic simply by

$$(R, \times, K) = (R, \times', K) = (R, \times'', K), \text{ etc.} \tag{4.6}$$

R is closed with respect to each of the operations $\times, \times', \times'', \cdots \xi, \rho', \rho'', \cdots$. A K-logical concept of a ring $(R, +, \times)$ is one which is definable entirely in terms of the K-logic of the ring. Among the K-logical concepts of a ring

are of course all multitations belonging to the class $|\times, K|$, i.e., all multitations compositionally generable from $\times$ and the $\rho \varepsilon K$. We note that

$$(4.7) \qquad |\times, K| = |\times', K| = \cdots = |\times, \times', \times'', \cdots, \xi, \rho', \rho'', \cdots|.$$

Further, if $\rho_1, \rho_2, \cdots$ are a set of generators of the group K one has

$$(4.8) \qquad |\times, K| = |\times, \rho_1, \rho_2, \cdots|.$$

In [1] and also in [6] special cases of the K-logical definability of the ring sum, $+$, — and hence of the whole ring, were treated. We here generalize and simultaneously clarify these applications by introducing several refinements. We say that a ring $(R, +, \times)$ is: (a) K-logically *definable* if its $+$ is a K-logical concept of the ring. (b) K-logically *equationally* definable if its $+$ is $\varepsilon\, |\times, K|$, i.e., if its $+$ satisfies an identity

$$(4.9) \qquad a + b \underset{a,b}{\equiv} \phi(a, b) \text{ where } \phi \,\varepsilon\, |\times, K|.$$

(c) K-logically *fixed*, if it is K-logically definable, and if there exists no other ring $(R, +_1, \times)$,—on the same set R and with the same product $\times$, but with $+_1 \neq +$, which is K-logically definable.

We shall illuminate these distinctions, and at the same time establish their essential independence, by again returning to the *simple* or C-logics (C = complementation group (2.1)).

A Boolean ring is C-logically equationally definable, since in such a ring (see [1] and 1),

$$(4.10) \qquad a + b = ab^* \otimes a^*b.$$

It is easy to extend this to the

THEOREM 4. *A Boolean ring is C-logically equationally definable and fixed.*

Proof. Let $(R, +, \times)$ be a Boolean ring, and let $(R, +_1, \times)$ be a ring having the same C-logic as $(R, +, \times)$. By Stone's theorem [8], a Boolean ring is characterized by the idempotency condition

$$(4.11) \qquad x \times x = x^2 = x \qquad (x \,\varepsilon\, R)$$

from which follows that

$$(4.12) \qquad x + x = 0.$$

Hence $(R, +_1, \times)$ is also Boolean, and

$$(4.13) \qquad x +_1 x = 0.$$

In addition, by hypothesis,

$$(4.14) \qquad 1 - x = 1 -_1 x = x^*.$$

Now from (4.11) and (4.12), respectively (4.13), follows

$$(4.15) \qquad x + y = x +_1 y \underset{x,y}{\equiv} xy^* \otimes x^*y,$$

which proves the theorem.

We know also that the more general class of *Boolean-like* rings (see [1]) are C-logically equationally definable, with $a + b$ again given by (4.10); in fact this was essentially the definition of this class of rings. In contrast to Theorem 4, however, we have

THEOREM 5. *A Boolean-like ring is C-logically equationally definable, but not in general C-logically fixed.*

Proof. Consider the two rings $(R, +, \times)$ and $(R, +_1, \times)$, where $R = \{0, 1, 2, 3\}$ and where

$$(4.16) \qquad \begin{array}{c|cccc} \times & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 0 & 2 & 0 & 2 \\ 3 & 0 & 3 & 2 & 1 \end{array} \qquad \begin{array}{c|cccc} + & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 1 & 2 & 3 \\ 1 & 1 & 0 & 3 & 2 \\ 2 & 2 & 3 & 0 & 1 \\ 3 & 3 & 2 & 1 & 0 \end{array} \qquad \begin{array}{c|cccc} +_1 & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 1 & 2 & 3 \\ 1 & 1 & 2 & 3 & 0 \\ 2 & 2 & 3 & 0 & 1 \\ 3 & 3 & 0 & 1 & 2 \end{array}.$$

Here $(R, +, \times) = H_4$, the simplest Boolean-like ring which is not also Boolean (see [1]), while $(R, +_1, \times) = ((4))$ is the ring of residues mod 4. From (4.16) it is easily found that whether * be computed from H_4 or from $((4))$, the results are identical,

$$(4.17) \qquad x^* = 1 - x = 1 -_1 x; \qquad * \frac{0\ 1\ 2\ 3}{1\ 0\ 3\ 2}.$$

Hence we have two distinct (even non-isomorphic) rings having identical C-logics, which proves Theorem 5.[6]

It was shown in [6] that a field $(F, +, \times)$ is always C-logically definable,—one such definition is recalled in (1.10) of **1**. While this particular C-logical definition of $+$ is obviously not equational, an equational one might conceivably exist. We show

[6] We have, of course, simultaneously shown that $((4))$ is not C-logically fixed.

THEOREM 6. (a) *A Field* $(F, +, \times)$ *is a C-logically fixed ring. However* (b) *in general a field will have no C-logical equational definition.*

Proof of (a): We first prove the

LEMMA. *Let* $(R, +, \times)$ *and* $(R, +_1, \times)$ *be rings, and let* $(R, +, \times)$ *possess no 0-divisors. Then* $(R, +_1, \times)$ *has no 0-divisors, and*

$$-x = -_1 x \qquad (x \varepsilon R). \tag{4.18}$$

Proof. Since $x(-1) = -x$, $x(-_1 1) = -_1 x$, it is obviously sufficient to show that (A): $-1 = -_1 1$. We have

$$\{(-1)(-_1 1)\}\{(-1)(-_1 1)\} = (-1)^2(-_1 1)^2 = 1, \tag{4.19}$$

since -1 commutes with all elements. Since no 0-divisors exist, $x^2 = 1 \rightarrow x = 1$ or $x =$ additive inverse of 1. Hence

$$(-1)(-_1 1) = 1 \text{ or } -1 \text{ or } -_1 1. \tag{4.20}$$

If $(-1)(-_1 1) = 1$, then $(-_1 1) = -1$, and (A) is proved. If $(-1)(-_1 1) \neq 1$, then since $\times$ is the same in both rings,

$$(-1)(-_1 1) = -1 = -_1 1. \tag{4.21}$$

Hence (A) and with it the Lemma is proved.

Note. A comparison of H_4 and ((4)) shows that the Lemma is in general false if 0-divisors are present.

The proof of part (a) of Theorem 6 is now immediate. Let $(F, +, \times)$ be a field and let $(F, +_1, \times)$ be a ring having the same C-logic as $(F, +, \times)$. We must show that $+ = +_1$. By hypothesis

$$1 - x = 1 -_1 x. \tag{4.22}$$

By the Lemma we then have

$$x + 1 = x +_1 1. \tag{4.23}$$

We must show that

$$x + y = x +_1 y. \tag{4.24}$$

For $x = 0$ this is trivial, and for $x \neq 0$, by (4.23),

$$x + y = x(1 + yx^{-1}) = x(1 +_1 yx^{-1}) = x +_1 y, \tag{4.25}$$

where x^{-1} is the $\times$ inverse of x. This proves part (a) of Theorem 6.

Proof of part (b). We shall show that $(F_3, +, \times)$, the field of residues mod 3, is not equationally C-logically definable. We recall a special instance of a well known theorem:

(B) Each of the 3^3 monotations $\alpha(x)$ of the class F_3 may be (uniquely) expressed in the 'analytical' form,

$$\alpha(x) = a_0 + a_1 x + a_2 x^2 \qquad (a_0, a_1, a_2 = 0, 1, 2 \pmod 3)). \tag{4.26}$$

We next show

(C) If $\phi(x)$ is a monotation of F_3 which is $\varepsilon \mid \times, C \mid$, then if $\phi(x)$ is expressed in the canonical form (4.26), $a_0 \neq 2$.

The truth of (C) is seen as follows. The class Γ consisting of all elements of $\mid \times, C \mid$ which are monotations, is inductively defined as follows: $1°$) $x \,\varepsilon\, \Gamma$; $2°$) if $\sigma(x)$ and $\tau(x)$ are ε's Γ, so are $\sigma^*(x)$, $\sigma^2(x)$, $\sigma(x) \times \tau(x)$. That we need go no higher than σ^2 follows from the identity

$$x^3 \equiv x \qquad (x \,\varepsilon\, F_3). \tag{4.27}$$

We see that x satisfies (C). Also, if $\sigma(x)$ and $\tau(x)$ each satisfy (C), so do σ^2, σ^* and $\sigma \times \tau$, the first and last since the constant term of the product $\sigma \times \tau$ is the product of the constant terms, and the second since $1^* = 0$, $0^* = 1$. Hence, by induction, (C) is proved.

We further observe that in F_3

(D) $$1 \equiv (x \times x^{2*})^* = \varepsilon\, \Gamma.$$

Suppose now that

$$x + y \text{ is } \varepsilon \mid \times, C \mid .$$

Then by (D), $1 + 1 = 2$ would be $\varepsilon \mid \times, C \mid$. This is however impossible by (C), which contradiction completes part (b), and with it Theorem 6.

By considering the prime subfield of a given field one can, with only minor modifications, strengthen part (b) of Theorem 6 to

THEOREM 7. *If F is a field of characteristic $\neq 2$, it cannot be equationally C-logically defined.*

Similarly one may show

THEOREM 8. *The ring $(W, +, \times)$ of whole numbers cannot be equationally C-logically defined.*

5. Ring-logics (K). Let $(R, +, \times)$ be an arbitrary but fixed ring, and K a group of permutations of R. We say that the group K is *semi-adapted* to the ring R if the ring is K-logically fixed; if in addition the ring is equationally definable in terms of its K-logic, we say that K is *fully-adapted* to the ring. We have just seen, for instance, that the simple complementation group C is always fully adapted to Boolean rings, but in general only semi-adapted to a given field.

If K is fully adapted to a ring R, we shall also refer to R as a *ring-logic* (K), or a *logic-ring* (K); in this case ring and K-logic uniquely and equationally fix each other, and it is therefore appropriate to speak of *the* ring of the K-logic, as well as the K-logic of the ring.

It is natural to inquire: Given a ring R, does there always exist at least one group K which is (a) semi-adapted, or (b) fully-adapted to R? Question (a) may be affirmatively answered by a simple construction into which we shall not here enter. Entirely different in nature is the stronger question (b), to which no complete answer has as yet been found. This latter question (b) may be restated: May any ring be converted into a ring-logic (K) by suitably choosing K?

That Boolean rings are not the only ring-logics we shall explicitly show in **8**. In general we may anticipate that the K-ality theory (see **2**) of a ring-logic (K) will be a combinatorially rich theory, in view of the unique equational determinancy of ring in terms of logic and conversely.

6. p-rings. In seeking interesting ring-logics other than Boolean rings we are led to a natural generalization of this latter class.[7] Let p be an arbitrary fixed prime integer. By a *p-ring* we mean a commutative ring with unit $(S, +, \times)$, in which, for all $a \varepsilon S$,

$$a^p = a \times a \times a \times \cdots \times a = a \tag{6.1}$$

$$p^a = a + a + a + \cdots + a = 0. \tag{6.2}$$

The class of Boolean rings is thus coextensive with the class of 2-rings ($p = 2$). For this special case, $p = 2$, (6.2) is a consequence of (6.1), as is well known. That this is not so in general is seen from ((6)), the ring of residues mod 6, in which (6.1) is satisfied with $p = 3$, but not (6.2). It is also easily shown that the prime p is unique for a given p-ring, i. e., that a ring cannot be both a p-ring and a p'-ring, with $p' \neq p$.

---

[7] As already noted, "p-rings" were first introduced in [9], where a proof of Theorem 9 is given.

It is not our purpose here to enter into the structure, either elementary or ideal, of p-rings; we mention however in passing the following extension of a familiar Boolean theorem as

THEOREM 9. *For given prime p, F_p = field of residues mod p is a p-ring, and is a sub-ring of any p-ring. If S is a finite p-ring,*

$$S = F_p \times F_p \times \cdots \times F_p, \tag{6.3}$$

where $\times$ denotes direct product; and hence the number of elements in a finite p-ring is always a power of p, p^t.

Let $S = (S, +, \times)$ be a p-ring. By the *cyclic* (*negation*)[8] group N of S we understand the group of coordinate transformations in S generated by $^\wedge$, where

$$a^\wedge = 1 + a. \tag{6.4}$$

(Here the order p of the cyclic group N is the (prime) characteristic p of N, unlike the complementation group C which always has the fixed order 2).

If $\times, \times', \times'', \cdots$ denote the transforms of $\times$ by $^\wedge$, respectively by $^{\wedge\wedge}$, etc., and if $+', +'', \cdots$ and $-', -'', \cdots$ have similar meanings, by (3.5), (6.4) one easily computes,

$$a \times^{(r)} b = a \times b + r(a + b) + r^2 - r. \tag{6.5}$$

$$a +^{(r)} b = a + b + r. \tag{6.6}$$

$$a -^{(r)} b = a - b - r. \tag{6.7}$$

Again, by Theorem 3, the operation $^\wedge$ is N-fixed, i. e., the same in each permissible coordinate system; this applies also to each of the operations $^{\wedge\wedge}, ^{\wedge\wedge\wedge}, \cdots$.

In a formula such as (6.4) the coefficient r (mod p) in $r(a + b)$ is an *apparent* (or removable) constant, since it represents $(a + b) + (a + b) + \cdots + (a + b)$. By contrast, the 'additive' constants, such as $r^2 - r$ (mod p) in (6.5), are *real* constants.

We now state, without proof, the p-ality theorem,—the generalization to p-rings of the classic Boolean duality. The proof of the theorem is much like that of the simple duality theory, and offers no difficulty.

THEOREM 10. *p-Ality Theorem for p-Rings. Let S be a p-ring, and let*

[8] The terminology 'cyclic negation' is borrowed from the expression by the same name in many-valued logics. See **8, 9**.

$$P(0, 1, 2, \cdots, p-1; \times, \times', \times'', \cdots; +, +', +'', \cdots;$$
$$-, -', -''; {}^\wedge; {}^{\wedge\wedge}; {}^{\wedge\wedge\wedge}; \cdots)$$

be any true proposition in S involving no apparent constants. Then each of the p-al propositions

$$P' = P(p-1, 0, 1, 2, \cdots; \times', \times'', \cdots; +', +'', \cdots;$$
$$-', -'', \cdots; {}^\wedge; {}^{\wedge\wedge}; {}^{\wedge\wedge\wedge}; \cdots)$$

$$P'' = P(p-2, p-1, 0, 1, 2, \cdots; \times'', \times''', \cdots; +'', +''', \cdots;$$
$$-'', -''', \cdots; {}^\wedge; {}^{\wedge\wedge}; {}^{\wedge\wedge\wedge}; \cdots), \text{ etc.}$$

obtained by (a) *leaving each of the operations* ${}^\wedge; {}^{\wedge\wedge}; \cdots$ *unchanged,* (b) *applying any cyclic permutation 'cogrediently' to all other operations and* (c) *the 'contragredient'* (= *inverse*) *permutation to the real constants, is again a true proposition of the ring S.*

The 'contragredient' element of the theorem is one not apparent in the simple case, $p = 2$, since in this case there is no difference between $x^\wedge$ and its inverse, $x^\wedge = x^* = x^\vee$. The self-p-ality of the 'cyclic negation' ${}^\wedge$, as well as that of ${}^{\wedge\wedge}$; ${}^{\wedge\wedge\wedge}; \cdots$ etc., is a consequence of Theorem 3.

7. 3-ring-logics. We shall now specialize to the case of 3-rings ($p = 3$). Here we explicitly show that the cyclic negation group N (of order 3) is fully adapted to this class of rings, and hence that each 3-ring is a ring-logic (N). The class of 3-ring-logics contains the 3-valued logic ($= F_3$) as its simplest representative, and the 3-valued logic and the general 3-ring-logics are related to each other and dominated by the *tri-ality theory,* exactly as are the ordinary logic of propositions ($= F_2$), 2-ring-logics (= Boolean rings = Boolean algebras) and the enfolding simple (Boolean) duality theory. We may, in this sense, speak of a unified theory. (See **8**, **9**).

Let S be a 3-ring. We have then

(7.1) $\quad a^\wedge = 1 + a, \qquad a^{\wedge\wedge} = 1 + 1 + a = 2 + a, \qquad a^{\wedge\wedge\wedge} = a.$

Let us abbreviate

(7.2) $\quad a^{\wedge\wedge} = a^\vee.$

We have

(7.3) $\quad a^\wedge = a^{\vee\vee}; \; a^{\wedge\wedge\wedge} = a^{\vee\vee\vee} = a^{\wedge\vee} = a^{\vee\wedge} = a;$

$$0^\wedge = 1, \quad 1^\wedge = 2, \quad 2^\wedge = 0; \qquad 0^\vee = 2, \quad 1^\vee = 0, \quad 2^\vee = 1$$

By definition of $\times'$, etc., we have

$$(7.4)\quad a \times' b = (a^{\wedge} \times b^{\wedge})^{\vee};\ a \times'' b = (a^{\vee} \times b^{\vee})^{\wedge};\ a +' b = (a^{\wedge} + b^{\wedge})^{\vee};$$
$$a +'' b = (a^{\vee} + b^{\vee})^{\wedge};\ a -' b = (a^{\wedge} - b^{\wedge})^{\vee};\ a -'' b = (a^{\vee} - b^{\vee})^{\wedge}.$$

In each of the relations (7.4) the $+$, $\times$ coordinate system is exhibited in a preferred role, in that each operation $\times', \times'', +', \cdots$ etc. is expressed in this coordinate system. This preferred role is removed, and each operation may then be expressed in any of the (three) permissible coordinate systems, by applying the *tri-ality* theorem, that is the p-ality theorem for $p = 3$. Each relation is then one of a *tri-al* set. From the first two relations (7.4), by tri-alization, we get for the tri-al ring products of a 3-ring, the

THEOREM 11. *Transformation Theorem* (= '*De Morgan' Formula for 3 Rings*). *In any 3-ring,*

$$(7.5)\quad \begin{aligned} a \times' b &= (a^{\wedge} \times b^{\wedge})^{\vee} = (a^{\vee} \times'' b^{\vee})^{\wedge} \\ a \times'' b &= (a^{\vee} \times b^{\vee})^{\wedge} = (a^{\wedge} \times' b^{\wedge})^{\vee} \\ a \times b &= (a^{\wedge} \times'' b^{\wedge})^{\vee} = (a^{\vee} \times' b^{\vee})^{\wedge}. \end{aligned}$$

(It may be noted that the *simple* ring 'De-Morgan' theorem (1.5) may be obtained by degeneration from (7.5) by taking $a^{\wedge} = a^{\vee} = a^*$ and writing $\times' = \otimes$).

From its derivation it is easily seen that Theorem 11 gives the correct formulas not only for converting the ring products from one coordinate system to another, but also for converting any multitation (= operation) $\phi(a, b, \cdots)$ in the 3-ring. Thus, if ϕ', ϕ'' are the transforms (3.5) of ϕ by $^{\wedge}$ and by $^{\vee}$ respectively, then, as in (7.5), we have

$$(7.6)\quad \begin{aligned} \phi'(a, b, \cdots) &= \{\phi(a^{\wedge}, b^{\wedge}, \cdots)\}^{\vee} = \{\phi''(a^{\vee}, b^{\vee}, \cdots)\}^{\wedge} \\ \phi''(a, b, \cdots) &= \{\phi(a^{\vee}, b^{\vee}, \cdots)\}^{\wedge} = \{\phi'(a^{\wedge}, b^{\wedge}, \cdots)\}^{\vee} \\ \phi(a, b, \cdots) &= \{\phi''(a^{\wedge}, b^{\wedge}, \cdots)\}^{\vee} = \{\phi'(a^{\vee}, b^{\vee}, \cdots)\}^{\wedge}. \end{aligned}$$

In the particular case of bitations we also write

$$\phi(a, b) = a\, \phi\, b$$

and the formulas (7.6) correspondingly, i.e.,

$$(7.7)\quad \begin{aligned} a\, \phi'\, b &= (a^{\wedge}\, \phi\, b^{\wedge})^{\vee} = (a^{\vee}\, \phi''\, b^{\vee})^{\wedge} \\ a\, \phi''\, b &= (a^{\vee}\, \phi\, b^{\vee})^{\wedge} = (a^{\wedge}\, \phi'\, b^{\wedge})^{\vee} \\ a\, \phi\, b &= (a^{\wedge}\, \phi''\, b^{\wedge})^{\vee} = (a^{\vee}\, \phi'\, b^{\vee})^{\wedge}. \end{aligned}$$

The six general transformation formulas (7.6), and hence in particular (7.7) and (7.4), may be condensed into a single convenient formula by means of a simple rule of thumb. Let us agree to call the 'cyclical negation' operation, $^\wedge$, the *predecessor*, and $^\vee$ ($= {}^{\wedge\wedge}$) the *successor* negation. Similarly, if k and κ are any two different members of the set $0, 1, 2$ we say that 'k is the *predecessor of* κ,'—also read, 'κ is the *successor* of k,' if $k \to \kappa$ in the 'standard' cyclic permutation (012). Since we have to do with a class of only three elements, we have the evident dichotomy: for any given k, κ with $k \neq \kappa$, of the two relations (1°) : k is the predecessor of κ, (2°) : κ is the predecessor of k, one and only one holds.

To *pre* a ring element, a, is to replace it by $a^\wedge$; to *pre* a ring operation is to replace it by its predecessor. A similar terminology holds with respect to *suc*. For instance, pre $\phi = \phi''$, suc $+ = +'$, suc $a = a + 2, \cdots$ etc. We may now reformulate the generalized Theorem 11, which we do in two slightly different ways (A) and ($\mathfrak{A}$),—the latter for more convenient application.

THEOREM 12. FUNDAMENTAL TRANSFORMATION THEOREM FOR MULTITATIONS. *In a 3-ring, let ϕ' and ϕ'' be the transposes (7.6) of a multitation ϕ by $^\wedge$ and by $^\vee$ respectively; let $\phi^{(k)}$, $\phi^{(\kappa)}$ be elements of the set ϕ, ϕ', ϕ''; let $^\oplus$, $^\ominus$ be some arrangement of the set $^\wedge$, $^\vee$.*

(A) *For given k and κ, with $k \neq \kappa$, the formula expressing $\phi^{(k)}$ in terms of $\phi^{(\kappa)}$ is given by*

$$\phi^{(k)}(a, b, \cdots) = \{\phi^{(\kappa)}(a^\oplus, b^\oplus, \cdots)\}^\ominus, \tag{7.8}$$

where $^\oplus$ must be chosen to 'agree with' κ, that is, $^\oplus$ is the predecessor negation, $^\wedge$, if κ is the predecessor of k, and is the successor negation, $^\vee$, if κ is the successor of k.

($\mathfrak{A}$) *For given k and given $^\oplus$,*

$$\{\phi^{(k)}(a, b, \cdots)\}^\oplus = \phi^{(\kappa)}(a^\oplus, b^\oplus, \cdots), \tag{7.9}$$

where κ must be chosen to 'agree with' $^\oplus$, in the above sense.

By repeated application of ($\mathfrak{A}$) of Theorem 12, and by use of the *pre*-ing and *suc*-ing terminology preceding Theorem 12, we have the very useful extension,

THEOREM 13. *Let Ψ be a multitation built up (by composition) from one or more 'component' multitations. Then a formula for $\Psi^\wedge$ is obtained*

by pre-ing everything in Ψ, that is, component multitations as well as ring element arguments. Similarly a formula for $\Psi^{\vee}$ is obtained by suc-ing everything in Ψ.

As illustrations of the application of Theorem 13, we have

$$
(7.10)\qquad \begin{aligned} &\{(a \times b) \times'' c\}^{\wedge} = (a^{\wedge} \times'' b^{\wedge}) \times' c^{\wedge} \\ &\{(2 +' a) \times'' (b^{\vee} \times' c)\}^{\vee} = (0 +'' a^{\vee}) \times (b^{\wedge} \times'' c^{\vee}). \end{aligned}
$$

Note. The treatment of 'constants' in Theorem 13 is exactly like that of a variable, unlike the 'contragredient' distinction required in the p-ality theorem.

Each of the relations (7.5), as well as future identities, may be directly verified by expressing each side in the same coordinate system by means of (7.1), (7.2) and the relations (6.4)-(6.6) written for $p = 3$, namely

$$
(7.11)\qquad \begin{aligned} &a \times' b = a \times b + a + b, \qquad a \times'' b = a \times b + 2(a + b) + 2, \\ &a +' b = a + b + 1, \qquad a +'' b = a + b + 2, \\ &a -' b = a - b - 1, \qquad a -'' b = a - b - 2. \end{aligned}
$$

Here again these exhibit the $+, \times$ coordinate system in a preferred role; this may be removed by tri-alizing the relations (7.11), whence each operation, in any coordinate system, is expressible in terms of operations belonging entirely to any given coordinate system:

THEOREM 14. *In a 3-ring,*

$$
(7.12)\qquad \begin{aligned} &a \times' b = a \times b + a + b = a \times'' b +'' 2(a +'' b) +'' 0 \\ &a \times'' b = a \times b + 2(a + b) + 2 = a \times' b +' a +' b \\ &a \times b = a \times'' b +'' a +'' b = a \times' b +' 2(a +' b) +' 1 \\ &a +' b = a + b + 1 = a +'' b +'' 0 \\ &a +'' b = a + b + 2 = a +' b +' 0 \\ &a + b = a +'' b +'' 2 = a +' b +' 1 \\ &a -' b = a - b - 1 = a -'' b -'' 0 \\ &a -'' b = a - b - 2 = a -' b -' 0 \\ &a - b = a -' b -' 1 = a -'' b -'' 2. \end{aligned}
$$

8. 3-ring-logics (continued). We now show that each 3-ring $(S, +, \times)$ is N-logically fixed, and moreover equationally. By the logic ($=$ logical algebra) of a 3-ring we shall here always understand the N-logic,

$$(S, \times, \times', \times'', {}^{\wedge}, {}^{\vee}). \tag{8.1}$$

THEOREM 15. *Each 3-ring $(S, +, \times)$ is a ring-logic, with $+$ logically definable by the equation*

$$a + b = ab^{\wedge} \times' a^{\wedge}b \times' a^2b^2. \tag{8.2}$$

Proof. The identity (8.2) may be verified by direct substitution from (7.1) and (7.12), making use of the 3-ring definitive properties (6.1) and (6.2) for $p = 3$. In the $+$, $\times$ coordinate system the right of (8.2) then becomes

$$ab^{\wedge} + a^{\wedge}b + ab^{\wedge}a^{\wedge}b + a^2b^2 + a^2b^2(ab^{\wedge} + a^{\wedge}b + ab^{\wedge}a^{\wedge}b), \tag{8.3}$$

which readily reduces to $a + b$. There remains to show that $(S, +, \times)$ is fixed by its logic. Suppose $(S, +_1, \times)$ is a ring having the same logic as $(S, +, \times)$. We must show that $+ = +_1$. By hypothesis

$$1 + a = 1 +_1 a, \tag{8.4}$$

from which one finds that $3a = 0$. Hence, since $\times$ is the same for both rings, $(S, +_1, \times)$ is also a 3-ring, by definition (6.1), (6.2). We may now re-verify (8.2) with $a +_1 b$ on the left, which shows that $+ = +_1$, and proves Theorem 15.

By tri-alizing (8.2), either directly as given or after expressing the right side in various 'pure' or 'mixed' ways by use of the transforming relations (7.12), one may obtain many similar formulas. We here note only one example. If we start with (8.2) expressed in a pure form, we get the tri-al set:

THEOREM 16. *In a 3-ring,*

$$a + b = \{(ab^{\wedge})^{\wedge}(a^{\wedge}b)^{\wedge}(a^2b^2)^{\wedge}\}^{\vee} \tag{8.5}$$

$$a +' b = \{(a \times' b^{\wedge})^{\wedge} \times' (a^{\wedge} \times' b)^{\wedge} \times' (a \times' a \times' b \times' b)^{\wedge}\}^{\vee} \tag{8.6}$$

$$a +'' b = \{(a \times'' b^{\wedge})^{\wedge} \times'' (a^{\wedge} \times'' b)^{\wedge} \times'' (a \times'' a \times'' b \times'' b)^{\wedge}\}^{\vee}. \tag{8.7}$$

Again, from each formula for $a +' b$, such as (8.6), and similarly for $a +'' b$, by recalling that $(a +' b)^{\vee} = a + b$ and $(a +'' b)^{\wedge} = a + b$, and by use of Theorem 13, we obtain numerous new formulas for $a + b$; we here mention one such, obtained from (8.6),

$$a + b = (a^{\vee} \times'' b)(a \times'' b^{\vee})(a^{\vee} \times'' a^{\vee} \times'' b^{\vee} \times'' b^{\vee}). \tag{8.8}$$

9. 3-valued logic. We here give a very brief orientation of 3-valued logic within the framework of general 3-ring-logics, and consider an illustration of the tri-ality theorem applied to the former. A more comprehensive treatment of 3-valued-logic, with the general p-ality theory as background, will be offered in a later communication.

Exactly as the logic of propositions ($=$ 2-valued logic) is mathematically equivalent to the simplest 2-ring ($=$ Boolean ring) $F_2 =$ ring ($=$ field) of 2 elements or 'truth values' 0, 1, so is the 3-valued logic equivalent to the simplest 3-ring $F_3 =$ ring ($=$ field) of 3 elements or 'truth values' 0, 1, 2.

By a well known theorem (holding as well in $F_p =$ field of residues mod $p =$ prime), (I): each multitation $\phi(x, y, \cdots)$ of the set F_3 may be 'analytically' expressed,—and moreover, uniquely, as a polynomial, mod 3, of the field $(F_3, +, \times)$. Thus the 3^3 monotations of the set F_3 are uniquely 'analytically' exhibited in the ring language by

$$\phi(x) = a + bx + cx^2, \tag{9.1}$$

and the $3^{(3^2)}$ bitations uniquely by

$$\phi(x, y) = a + bx + cy + dxy + ex^2 + fy^2 + gxy^2 + hx^2y + kx^2y^2; \tag{9.2}$$

etc.

The 'logical language' is concerned entirely with

$$(F_3, \times, \times', \times'', {}^\wedge, {}^\vee), \tag{9.3}$$

and the 'completeness' of this logic follows from (I) and Theorem 15, making it possible to formulate all possible multitations of F_3, i. e., all possible 3-valued propositional functions, entirely within the logic (9. 3).

Moreover, since in F_3,

$$| \times, {}^\wedge | = | \times, {}^\vee | = | \times', {}^\wedge | = \cdots = | \times, \times', \times'', {}^\wedge, {}^\vee |, \tag{9.4}$$

(see (4. 7) and (4. 8)), in a formal development of this 3-valued logic of propositions it is sufficient to take only the operations $\times, {}^\wedge$, or else only $\times, {}^\vee$, etc., as undefined.

We consider an illustration of tri-al propositions. Let us read

$$0 = \text{false}, \qquad 1 = \text{true}, \qquad 2 = \text{indeterminate}. \tag{9.5}$$

Let $a \& b$ denote the propositional function which is true only if a and b are both true, and false otherwise. In the ring language (9. 2) we find (uniquely),

$$(9.6)\qquad a \,\&\, b = ab + ab^2 + a^2b + a^2b^2,$$

and in the logical language

$$(9.7)\qquad a \,\&\, b = ab(a \times' b)^{\wedge} = aa^{\wedge}bb^{\wedge}.$$

The tri-als of (9.7) are

$$(9.8)\qquad \begin{aligned} a \,\&'\, b &= a \times' b \times' (a \times'' b) = a \times' a^{\wedge} \times' b \times' b^{\wedge} \\ a \,\&''\, b &= a \times'' b \times'' (ab)^{\wedge} = a \times'' a^{\wedge} \times'' b \times'' b^{\wedge}. \end{aligned}$$

Here, in words,

(9.9)
$a \,\&'\, b$ is false only if a and b are both false, and indeterminate otherwise.
$a \,\&''\, b$ is indeterminate only if a and b are both true, and true otherwise.

10. Conjectures, problems. We have seen that the cyclic negation group N is fully adapted to both 2-rings (Boolean rings), and 3-rings; otherwise stated, that a p-ring is a ring-logic (N) for $p=2$ and $p=3$. Is this true for all primes p, or indeed for *some* prime $p > 3$? This question still remains unanswered. It may be shown, exactly as for 3-rings, that a p-ring is N-logically fixed if it is N-logically equationally definable. Hence the above question has an affirmative answer for such and only such primes p for which an identity similar to (8.2) exists.

In the case of a Boolean ring (2-ring) the group N reduces to the complementation group C, and hence in a Boolean ring $* = {}^{\wedge}$, $\otimes = \times'$, $\oplus = +'$, etc. (see **1**). It is instructive to compare the logical formulas for $+$ given in (1.8) and (8.2) for $p=2$ and for $p=3$,

$$(10.1)\qquad a + b = ab^{\wedge} \times' a^{\wedge}b \quad \text{(in a 2-ring)}$$

$$(10.2)\qquad a + b = ab^{\wedge} \times' a^{\wedge}b \times' a^2b^2 \quad \text{(in a 3-ring)}.$$

It is easily checked that the 3-case does not 'cover' the 2-case, i.e., that the formula (10.2) does not give the correct definition for $+$ when it is applied to a 2-ring; similarly the 2-case does not cover the 3-case. There are of course other logical equational formulas for $+$, such as (8.5) and (8.8), and others,—both for 2- and for 3-rings. It is however to be conjectured that there exists no single formula which covers both cases, in the above sense; and similarly for any primes p, p' ($p' \neq p$) for which p-ring and p'-ring are both ring logics.

In this connection the formulas (10.1) and (10.2) suggest that a similar formula with 5 factors might exist for 5-rings, of the form

$$a + b = ab^{\wedge} \times' a^{\wedge}b \times' W \times' Y \times' Z. \tag{10.3}$$

It may, however, be shown that this is impossible, in fact that no formula of the type (10.3) exists which contains a 'factor' $ab^{\wedge} \times' a^{\wedge}b$.[9]

UNIVERSITY OF CALIFORNIA,
BERKELEY.

BIBLIOGRAPHY.

1. A. L. Foster, "The theory of Boolean-like rings," *Transactions of the American Mathematical Society*, vol. 59 (1946), pp. 166-187.
2. ———, "The idempotent elements of a commutative ring form a Boolean algebra; ring duality and transformation theory," *Duke Mathematical Journal*, vol. 12 (1945), pp. 143-152.
3. ———, "Maximal idempotent sets in a ring with unit," *Duke Mathematical Journal*, vol. 13 (1946), pp. 247-258.
4. ———, "On the permutational representation of general sets of operations by partition lattices," *Transactions of the American Mathematical Society*, vol. 66 (1949), pp. 366-388.
5. A. L. Foster and B. A. Bernstein, "Symmetric approach to commutative rings with duality theorem: Boolean duality as special case," *Duke Mathematical Journal*, vol. 11 (1944), pp. 603-616.
6. ———, "A dual symmetric definition of field," *American Journal of Mathematics*, vol. 67 (1945), pp. 329-349.
7. F. Harary, "The structure of Boolean-like rings," in process of publication.
8. M. H. Stone, "The theory of representations of Boolean algebras," *Transactions of the American Mathematical Society*, vol. 40 (1936), pp. 37-111.
9. N. H. McCoy and Deane Montgomery, "A representation of generalized Boolean rings," *Duke Mathematical Journal*, vol. 3 (1937), pp. 455-459.

[9] (Added in proof). Since this article was presented, the author has succeeded in confirming the above conjecture; it is shown that all p-rings are ring-logics (mod N). As might be expected for $p > 3$ the logical definition of $+$, corresponding to the special cases (10.1) and (10.2), is quite complicated. This result, which is to appear in the *University of California Publications in Mathematics*, is based on a comprehensive study of the structure of p-rings, in process of publication in *Acta Mathematica*.

ON LINEAR DIFFERENCE EQUATIONS OF SECOND ORDER.*

By Philip Hartman and Aurel Wintner.

The theorems to be proved below represent extensions to the case of difference equations of certain results proved in [2] for the case of differential equations. The standard proof of the theorem of Kneser, used *loc. cit.*, is not now available and will have to be replaced by another approach. The latter will be patterned after the method of [3]. It turns out that the resulting criterion is by necessity different from that prevailing in the case of differential equations.

The first of the theorems to be proved is as follows:

(I) *Let* $q_0, q_1, \cdots ; r_0, r_1, \cdots$ *be two sequences of real numbers satisfying the inequalities*

$$(1) \qquad 1 - r_k - q_k > 0 \text{ and } q_k > 0 \qquad (k = 0, 1, \cdots).$$

Then the difference equation

$$(2) \qquad \Delta^2 y_k + r_k \Delta y_k - q_k y_k = 0 \qquad (k = 0, 1, \cdots)$$

possesses a solution $y_0, y_1, \cdots$ *satisfying*

$$(3) \qquad y_k > 0 \text{ and } \Delta y_k > 0 \qquad (k = 0, 1, \cdots).$$

It is understood that $\Delta y_k = y_{k+1} - y_k$ and $\Delta^2 y_k = \Delta(\Delta y_k) = y_{k+2} - 2y_{k+1} + y_k$.

Kneser's theorem, which deals with the differential equation

$$(4) \qquad y'' - q(x) y = 0,$$

was extended in [2] to the equation

$$(5) \qquad y'' + r(x) y' - q(x) y = 0.$$

In the case of differential equations, it is supposed that $r(x)$, $q(x)$ are continuous for large x, and that $q(x) \geqq 0$, but no restriction is placed on $r(x)$. If $r(x) \equiv 0$, the analogue of conditions (1) becomes

$$1 - q(x) > 0 \text{ and } q(x) > 0.$$

* Received March 5, 1949.

The first of the latter two inequalities is not needed in the theorem on differential equations. On the other hand, (I) becomes false if the first condition of (1) is omitted. In fact, the constant 1 occurring in this condition is the best constant; in the sense that if the first inequality of (1) is replaced by $1-\epsilon-r_k-q_k>0$, where $\epsilon>0$, then (2) need not have a solution satisfying (3). This is illustrated by the equation $\Delta^2 y_k-(1+\epsilon)y_k=0$, where $k=0,1,2,\cdots$. In fact, every solution $y_0,y_1,\cdots$ of this equation is a linear combination of the solutions given by $y_k=(1\pm(1+\epsilon)^{\frac{1}{2}})^k$, where $k=0,1,\cdots$, but no linear combination can satisfy (3) when $\epsilon>0$.

If $1-r_k-q_k>0$ is weakened to $1-r_k-q_k\geqq 0$, then the assertion (3) must be weakened to allow $y_k=0$ for all large k. This is illustrated by case $\epsilon=0$ of the equation just considered; actually, the equation then reduces to the first order difference equation $y_{k+2}-2y_{k+1}=0$, where $k=0,1,\cdots$, which possesses $y_0=\text{const.}\geqq 0$, $y_1=y_2=\cdots=0$ as the only non-decreasing solution.

Corresponding to the situation in differential equations, the second condition in (1) can be relaxed to $q_k\geqq 0$, provided that q_k does not vanish for all large k. In the latter case, the second assertion in (2) must be relaxed to $\Delta y_k=0$ for large k.

It is known (Sturm; cf. [1], pp. 176-177) that the first inequality in (1) assures that the "zeros" of two non-trivial solutions of (2) separate each other. A "zero" is meant in the following sense: If $y_0,y_1,\cdots$ is a solution of (2), and if one considers the polygonal path joining the points (n,y_n) in the (x,y)-plane, then a point common to this path and to the x-axis is called a zero of the solution $y_0,y_1,\cdots$.

It will be shown that a non-trivial solution of (2) has at most one zero in virtue of (1). In view of the separation theorem, it is sufficient to exhibit a solution of (1) which has no zero. Such a solution is obtained by assigning, for example, the initial conditions $y_0=1$, $\Delta y_0=y_1-y_0=1$; the solution $y_0,y_1,\cdots$ is then determined uniquely, since (2) can be written in the form

$$y_{k+2}-(2p_k-r_k)y_{k+1}+(1-r_k-q_k)y_k=0 \qquad (k=0,1,\cdots).$$

That the solution, determined by the assigned initial conditions, has no zero is a consequence of the following assertion:

If $y_0,y_1,\cdots$ is a solution of (2) and if, for some fixed $n\geqq 0$,

$$y_n\geqq 0 \text{ and } \Delta y_n\geqq 0, \tag{6}$$

then

$$y_k\geqq y_n;\ \text{in fact, } \Delta y_k\geqq 0 \qquad (k=n,n+1,\cdots). \tag{7}$$

In order to see this, rewrite the case $k=n$ of (2) in the form

$$\Delta y_{n+1}+(r_n-1)\Delta y_n=qy_n\geqq 0.$$

The first condition of (1) implies that $1-r_n>q_n\geqq 0$. Consequently, by the last formula line, $\Delta y_{n+1}\geqq 0$. Hence, (7) holds for $k=n+1$ and, by induction, for all $k\geqq n$.

Since a non-trivial solution of (2) has at most one zero, it follows that if n, m is a given pair of integers satisfying $0\leqq m<n$, and if y_m, y_n is a given pair of numbers, then there exists one and only one solution $y_0, y_1, \cdots$ for which y_n, y_m assume the given values. In fact, every solution of (2) is determined by its pair of initial values y_0, y_1; hence, every solution $y_0, y_1, \cdots$ is a linear combination,

$$y_k=c_1y^1{}_k+c_2y^2{}_k, \tag{8}$$

of the pair of solutions $y^1{}_0, y^1{}_1, \cdots, y^2{}_0, y^2{}_1, \cdots$ determined by $y^1{}_0=1$, $y^1{}_1=0$ and $y^2{}_0=0$, $y^2{}_1=1$, respectively. Conversely, every linear combination (8) is a solution of (2). Thus, what is required for the existence and the uniqueness of a solution, for which y_n, y_m are prescribed, is the unique solvability for c_1, c_2 of the linear equations

$$y_m=c_1y^1{}_m+c_2y^2{}_m \text{ and } y_m=c_1y^1{}_m+c_2y^2{}_m.$$

If the latter do not have a unique solution c_1, c_2, then there exists a pair of constants c_1, c_2, not both zero, such that

$$c_1y^1{}_m+c_2y^2{}_m=0 \text{ and } c_1y^1{}_n+c_2y^2{}_n=0.$$

But then the corresponding non-trivial solution (8) has two zeros, which is impossible. Consequently, if $m\neq n$, the numbers y_m, y_n determine a unique solution of (2).

In order to complete the proof of (I), let $y_0{}^j, y_1{}^j, y_2{}^j, \cdots$ denote the unique solution of (2) satisfying

$$y_0{}^j=1 \text{ and } y_j{}^j=0, \qquad (j=1,2,\cdots). \tag{9}$$

Then $y_k{}^j>0$ for $k=0,1,\cdots,j-1$, since otherwise the solution $y_0{}^j, y_1{}^j, \cdots$ had two zeros. Also

$$1=y_0{}^j>y_1{}^j>\cdots>y_j{}^j=0. \tag{10}$$

For, if $y_n{}^j\geqq y_{n+1}{}^j$, that is, if $\Delta y_n{}^j\geqq 0$ holds for some value of n on the interval $0\leqq n\leqq j-1$, then $y_k{}^j\geqq y_n{}^j>0$ for $k\geqq n$, since (6) implies (7). But this contradicts $y_j{}^j=0$. Consequently, (10) holds.

A diagonal selection process shows that the sequence of integers $j = 1, 2, \cdots$, contains a subsequence having the property that, if the j-th element of the subsequence is denoted simply by j, then, as $j \to \infty$, the limit $y_k = \lim y_k^j$ exists for $k = 0, 1, \cdots$. Clearly, this limit sequence $y_0, y_1, \cdots$ is a solution of (2) and satisfies

$$(11) \qquad y_k \geqq 0 \text{ and } \Delta y_k \leqq 0, \qquad (k = 0, 1, \cdots)$$

in virtue of (10). Furthermore, $y_0 = 1$, by (9); so that the solution $y_0, y_1, \cdots$ is not identically zero. Since it has at most one zero, it follows, from (11), that it has no zero, that is, that $y_k > 0$.

It remains to show that $\Delta y_n = 0$ cannot hold for any n. If this did occur for some n, then $\Delta y_k = 0$ for $k \geqq n$, by (11), since (6) implies (7). Hence, $\Delta^2 y_k = 0$ for $k \geqq n$; so that $q_k y_k = 0$ for $k \geqq n$, by (2). This implies $q_k = 0$ for $k \geqq n$, which contradicts the second assumption in (1).

The proof of (I) is now complete.

(II) *Let* $p_0, p_1, \cdots$; $q_0, q_1, \cdots$; $r_0, r_1, \cdots$ *be three sequences of numbers satisfying*

$$(12) \qquad p_k > 0; \; p_k - r_k - q_k > 0; \; q_k > 0, \qquad (k = 0, 1, \cdots),$$

and let the three sequences $\Delta p_0, \Delta p_1, \cdots$; $q_0, q_1, \cdots$; $r_0, r_1, \cdots$ *be completely monotone, that is, let*

$$(13) \qquad (-1)^n \Delta^{n+1} p_k \geqq 0; \qquad (-1)^n \Delta^n q_k \geqq 0; \qquad (-1)^n \Delta^n r_k \geqq 0,$$

where $k, n = 0, 1, 2, \cdots$. *Then the difference equation*

$$(14) \qquad p_k \Delta^2 y_k + r_k \Delta y_k - q_k y_k \qquad (k = 0, 1, \cdots)$$

possesses a positive, completely monotone solution $y_0, y_1, \cdots$:

$$(15) \qquad y_k > 0 \text{ and } (-1)^n \Delta^n y_k \geqq 0 \qquad (k, n = 0, 1, \cdots).$$

It is understood that

$$\Delta^n y_k = \Delta(\Delta^{n-1} y_k) = \sum_{m=0}^{n} C_m{}^n (-1)^{n-m} y_{m+k},$$

where the $C_m{}^n$ denote the binomial coefficients. The theorem (II) is an analogue of a theorem on differential equations proved in [2].

The proof of (II) proceeds as follows: If (14) is divided by p_k, which is permissible in view of the first assumption in (12), then the difference equation (14) is reduced to one of the form (2). Furthermore, the last two

conditions of (12) imply the conditions (1). Hence, (14) possesses a solution $y_0, y_1, \cdots$ satisfying (3). This means that

$$(16_n) \qquad (-1)^n \Delta^n y_k \geqq 0 \qquad (k = 0, 1, \cdots)$$

holds for $n = 0, 1$. In order to prove that (16_n) holds for every n, suppose that it holds for $n = 0, 1, \cdots, j + 1$.

If $j \geqq 0$ is fixed and if $a_0, a_1, \cdots; b_0, b_1, \cdots$ are given sequences, then

$$\Delta^j (a_k b_k) = \sum_{m=0}^{j} C_m{}^j (\Delta^{j-m} b_{k+m}) (\Delta^m a_k), \qquad (k = 0, 1, \cdots).$$

Thus, if the operator Δ^j is applied to (14) and the resulting equation is solved for $\Delta^{j+2} y_k$, it is seen that $p_{k+j} \Delta^{j+2} y_k$ equals

$$-\sum_{m=0}^{j-1} C_m{}^j (\Delta^{j-m} p_{k+m})(\Delta^{m+2} y_k) - \sum_{m=0}^{j} C_m{}^j (\Delta^{j-m} r_{k+m})(\Delta^{m+1} y_k) + \sum_{m=0}^{j} C_m{}^j (\Delta^{j-m} q_{k+m})(\Delta^m y_k).$$

The induction hypothesis and (13) show that every term in these sums vanishes or has the same sign as $(-1)^{j+2}$; hence,

$$(-1)^{j+2} p_{k+j} \Delta^{j+2} y_k \geqq 0 \qquad (k = 0, 1, \cdots).$$

Since $p_{k+j} > 0$, the induction and the proof of (II) are now complete.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] M. Bôcher, "Boundary problems in one dimension," *5th International Congress of Mathematicians, Proceedings*, vol. I (1912), pp. 163-195.

[2] P. Hartman and A. Wintner, "On the Laplace-Fourier transcendents," *American Journal of Mathematics*, vol. 71 (1949), pp. 367-372.

[3] A. Wintner, "On linear repulsive forces," *ibid.*, pp. 362-366.

ON THE UNIFORM CESÀRO SUMMABILITY OF CERTAIN SPECIAL TRIGONOMETRICAL SERIES.*

By CHING-TSÜN LOO.

1. It is well-known that if $\lambda_\nu \to 0, \Delta\lambda_\nu \geqq 0$,[1] then both series

$$(1.1) \qquad \tfrac{1}{2}\lambda_0 + \sum_{\nu=1}^{\infty} \lambda_\nu \cos \nu\theta, \qquad \sum_{\nu=1}^{\infty} \lambda_\nu \sin \nu\theta$$

are uniformly convergent in $(\epsilon, \pi - \epsilon)$, $\epsilon > 0$. We also know that if $\lambda_\nu \to 0$, and $\Delta^2\lambda_\nu \geqq 0$, then the first derived series of (1.1) are uniformly summable $(C, 1)$ in that interval.[2] It is interesting to see whether these theorems can be extended to a theorem of general scale. The purpose of this paper is to give a positive answer to this question.

THEOREM. *If $\lambda_\nu \to 0$ and $\Delta^{k+1}\lambda_\nu \geqq 0$, then the k-th derived series of (1.1) are uniformly summable (C, k) in any interval $(\epsilon, \pi - \epsilon)$, $\epsilon > 0$, where k is any integer $\geqq 0$.*

We write

$$(1.2) \qquad C_\nu^{(k)} = \lambda_\nu h_\nu^{(k)}, \qquad h_\nu^{(k)} = \nu^k A_{n-\nu}^{(k)}$$

where $A_\nu^{(k)} = \binom{k+\nu}{\nu}$ are the Cesàro numbers. We have to prove that

$$(1.3) \qquad \sigma_n^{(k)}(\theta) = \sum_{\nu=0}^{n} C_\nu^{(k)} e^{i\nu\theta} / A_n^{(k)}$$

is uniformly convergent in the interval $(\epsilon, \pi - \epsilon)$ as $n \to \infty$. Our theorem then follows by considering the semi-sums and semi-differences of $\sigma_n^{(k)}(\theta)$ and $\sigma_n^{(k)}(-\theta)$.

2. We shall first establish the following formula: If $p \geqq 1$, then

$$(2.1) \qquad \begin{aligned} A_n^{(k)}\sigma_n^{(k)}(\theta) = & \sum_{j=0}^{p-1} (-1)^j e^{ij\theta} \Delta^j C_0^{(k)} / (1 - e^{i\theta})^{j+1} \\ & + (-1)^p e^{ip\theta} / (1 - e^{i\theta})^p \sum_{\nu=0}^{n-p} e^{i\nu\theta} \Delta^p C_\nu^{(k)} \\ & + e^{i(n+1)\theta} \sum_{j=0}^{p-1} (-1)^{j+1} / (1 - e^{i\theta})^{j+1} \Delta^j C_{n-j}^{(k)}. \end{aligned}$$

* Received March 21, 1949.

[1] We write $\Delta^0\lambda_\nu = \lambda_\nu$, $\Delta^1\lambda_\nu = \Delta\lambda_\nu = \lambda_\nu - \lambda_{\nu+1}$ and $\Delta^k\lambda_\nu = \Delta(\Delta^{k-1}\lambda_\nu)$ for $k \geqq 1$.

[2] A. Zygmund, *Trigonometrical series* (Warszawa-Lwów, 1935), p. 129, Ex. 6.

Let

$$S_\nu = 1 + e^{i\theta} + \cdots + e^{i\nu\theta} = (1 - e^{i(\nu+1)\theta})/(1 - e^{i\theta}),$$

then

$$S_\nu - S_{\nu-1} = -(e^{i(\nu+1)\theta} - e^{i\nu\theta})/(1 - e^{i\theta})$$

and

$$(2.2) \qquad \sum_{\nu=0}^{n} C_\nu^{(k)} e^{i\nu\theta} = \sum_{\nu=0}^{n} C_\nu^{(k)} (S_\nu - S_{\nu-1})$$

$$= -\sum_{\nu=0}^{n} C_\nu^{(k)} (e^{i(\nu+1)\theta} - e^{i\nu\theta})/(1 - e^{i\theta})$$

$$= \{C_0^{(k)} - e^{i\theta} \sum_{\nu=0}^{n-1} e^{i\nu\theta} \Delta C_\nu^{(k)} - e^{i(n+1)\theta} C_n^{(k)}\}/(1 - e^{i\theta}),$$

where $C_0^{(k)} = 0$. Hence the formula (2.1) has been proved for $p = 1$.

The same device gives

$$(2.3) \qquad \sum_{\nu=0}^{n-(p-1)} e^{i\nu\theta} \Delta^{(p-1)} C_\nu^{(k)} = \{\Delta^{(p-1)} C_0^{(k)} - e^{i\theta} \sum_{\nu=0}^{n-p} e^{i\nu\theta} \Delta^p C_\nu^{(k)}$$

$$- e^{i(n-p+2)} \Delta^{(p-1)} C_{n-p+1}^{(k)}\}/(1 - e^{i\theta}).$$

Suppose (2.1) holds for $p - 1$. Replacing the middle term of the right side of (2.1) by left side of (2.3) and rearranging the corresponding terms, we get (2.1) for p.

3. Let us put $k + 1$ for p in (2.1). Let $S_{n\,1}(\theta)$, $S_{n\,2}(\theta)$ and $S_{n\,3}(\theta)$ be the first, second and third sums of the right side of that formula. We are going to prove that $S_{n\,3}(\theta)/A_n^{(k)} = o(1)$ and that each of the sums $S_{n\,1}(\theta)/A_n^{(k)}$ and $S_{n\,2}(\theta)/A_n^{(k)}$ tends to a finite limit uniformly with respect to θ in $(\epsilon, \pi - \epsilon)$.

In order to estimate the orders of $\Delta^l C_\nu^{(k)}$ for $0 \leq l \leq k + 1$, we use the formula

$$(3.1) \qquad \Delta^l \xi_\nu \eta_\nu = \sum_{q=0}^{l} C_{l\,q} \Delta^q \xi_\nu \Delta^{l-q} \eta_{\nu+q},$$

where C_{lq} are constants. First of all we have

$$(3.2) \qquad \Delta^l h_\nu^{(k)} = \Delta^l \nu^k A_{n-\nu}^{(k)} = \sum_{q=0}^{l} C_{l\,q} \Delta^q A_{n-\nu}^{(k)} \Delta^{l-q} (\nu + q)^k$$

$$= \sum_{q=0}^{l} C_{l\,q} A_{n-\nu}^{(k-q)} \Delta^{l-q} (\nu + q)^k,$$

since $\Delta A_{n-\nu}^{(k)} = A_{n-\nu}^{(k)} - A_{n-(\nu+1)}^{(k)} = A_{n-\nu}^{(k-1)}$, and in general $\Delta^l A_{n-\nu}^{(k)} = A_{n-\nu}^{(k-l)}$ for $0 \leq l \leq k$, and is identically zero for $l > k$. Next we use (3.1) again,

$$(3.3)\qquad \Delta^l C_\nu^{(k)} = \Delta^l \lambda_\nu h_\nu^{(k)} = \sum_{p=0}^{l} C_{l\,p} \Delta^p h_\nu^{(k)} \Delta^{l-p} \lambda_{\nu+p}$$
$$= \sum_{p=0}^{l} C_{l\,p} \sum_{q=0}^{p} C_{p\,q} A_{n-\nu}^{(k-q)} \Delta^{p-q} (\nu+q)^k \Delta^{l-p} \lambda_{\nu+p}.$$

Observe that the terms in $S_{n\,3}(\theta)/A_n^{(k)}$ apart from corresponding factors $(-1)^{l+1} e^{i(n+1)\theta}/(1-e^{i\theta})^{l+1}$ are each of the form $\Delta^l C_{n-l}^{(k)}/A_n^{(k)}$, $l=0,1,\cdots,k$. Using (3.3), we obtain

$$\Delta^l C_{n-l}^{(k)}/A_n^{(k)} = O\Big(\sum_{p=0}^{l} \sum_{q=0}^{p} n^{k-p+q} \Delta^{l-p} \lambda_{n-l+p}/n^k\Big)$$
$$= O\Big(\sum_{p=0}^{l} \sum_{q=0}^{p} n^{-(p-q)} \Delta^{(l-p)} \lambda_{n-l+p}\Big) = o(1),$$

since $l \geq p \geq q$, and $\Delta^{l-p}\lambda_{n-l+p} = o(1)$. Thus we have

$$(3.4)\qquad S_{n\,3}/A_n^{(k)} = e^{i(n+1)\theta}/A_n^{(k)} \sum_{j=0}^{k} (-1)^{j+1}/(1-e^{i\theta})^{j+1} \Delta^j C_{n-j}^{(k)} = o(1)$$

uniformly in $(\epsilon, \pi-\epsilon)$.

Next observe that terms in $S_{n\,1}(\theta)/A_n^{(k)}$ apart from corresponding factors $(-1)^l e^{il\theta}/(1-e^{i\theta})$ are each of the form $\Delta^l C_\nu^{(k)}/A_n^{(k)}$, $l=0,1,\cdots,k$. Since we have

$$\Delta^l C_0^{(k)}/A_n^{(k)} = 1/A_n^{(k)} \sum_{q=0}^{l} (-1)^q \tbinom{l}{q} A_{n-q}^{(k)} q^k \lambda_q$$
$$= \sum_{q=0}^{l} (-1)^q \tbinom{l}{q} A_{n-q}^{(k)} q^k \lambda_q / A_n^{(k)}$$

which tends to $\sum_{q=0}^{l} (-1)^q \binom{l}{q} q^k \lambda_q$ as $n \to \infty$, it follows that

$$(3.5)\qquad S_{n\,1}(\theta)/A_n^{(k)} = 1/A_n^{(k)} \sum_{j=0}^{k} (-1)^j e^{ij\theta} \Delta^j C_0^{(k)}/(1-e^{i\theta})^{j+1}$$
$$\to \sum_{j=0}^{k} \sum_{q=0}^{j} (-1)^{j+q} \tbinom{j}{q} e^{ij\theta} q^k \lambda_q/(1-e^{i\theta})^{j+1}.$$

4. It remains to prove that

$$(4.1)\qquad S_{n\,2}(\theta)/A_n^{(k)} = \{(-1)^{k+1} e^{i(k+1)\theta}/(1-e^{i\theta})^{k+1} A_n^{(k)}\} \sum_{\nu=0}^{n-(k+1)} e^{i\nu\theta} \Delta^{k+1} C_\nu^{(k)}$$

tends uniformly to a limit.

To make the situation clear, we shall separate $\Delta^{k+1} C_\nu^{(k)}/A_n^{(k)}$ into three sums. By (3.2),

$$\Delta^{k+1} h_\nu^{(k)} = \sum_{q=1}^{k} C_{k+1\,q} A_{n-\nu}^{(k-q)} \Delta^{k+1-q} (\nu+q)^k$$

since $\Delta^{k+1}A_{n-\nu}^{(k)}=0$, $\Delta^{k+1}(\nu+q)^k=0$. By (3.3),

$$\Delta^{k+1}C_\nu^{(k)}/A_n^{(k)}=1/A_n^{(k)}\sum_{p=0}^{k}C_{k+1\,p}\sum_{q=0}^{p}C_{p\,q}A_{n-\nu}^{(k-q)}\Delta^{p-q}(\nu+q)^k\Delta^{k+1-p}\lambda_{\nu+p}$$

$$+C_{k+1\,k+1}/A_n^{(k)}\sum_{q=1}^{k}C_{k+1\,q}A_{n-\nu}^{(k-q)}\Delta^{k+1-q}(\nu+q)^k\lambda_{\nu+k+1}$$

$$=1/A_n^{(k)}\sum_{p=0}^{k}C_{k+1\,p}C_{p\,0}A_{n-\nu}^{(k)}\Delta^p\nu^k\Delta^{k+1-p}\lambda_{\nu+p}$$

$$+1/A_n^{(k)}\sum_{p=1}^{k}C_{k+1\,p}\sum_{q=1}^{p}C_{p\,q}A_{n-\nu}^{(k-q)}\Delta^{p-q}(\nu+q)^k\Delta^{k+1-p}\lambda_{\nu+p}$$

$$+C_{k+1\,k+1}/A_n^{(k)}\sum_{q=1}^{k}C_{k+1\,q}A_{n-\nu}^{(k-q)}\Delta^{k+1-q}(\nu+q)^k\lambda_{\nu+k+1}$$

$=J_{n\,\nu\,1}+J_{n\,\nu\,2}+J_{n\,\nu\,3}$, say. Putting

$$(4.2)\qquad 1/A_n^{(k)}\sum_{\nu=0}^{n-(k+1)}e^{i\nu\theta}\Delta^{k+1}C_\nu^{(k)}=\sum_{\nu=0}^{n-(k+1)}J_{n\,\nu\,1}e^{i\nu\theta}+\sum_{\nu=0}^{n-(k+1)}J_{n\,\nu\,2}e^{i\nu\theta}+\sum_{\nu=0}^{n-(k+1)}J_{n\,\nu\,3}e^{i\nu\theta}$$

$=W_{n\,1}(\theta)+W_{n\,2}(\theta)+W_{n\,3}(\theta)$, we are going to prove that $W_{n\,2}(\theta)=o(1)$, $W_{n\,3}(\theta)=o(1)$ uniformly in θ, and that $W_{n\,1}(\theta)$ tends uniformly to a limit. Before doing so, we shall first deduce from our assumptions $\lambda_\nu\to 0$, $\Delta^{k+1}\lambda_\nu\geq 0$, some simple consequences which are useful in the following proofs.

We observe that the assumptions imply that

$$(4.3)\qquad \begin{array}{ll}\text{(i)}\quad \Delta^l\lambda_\nu\geq 0, & \text{(ii)}\quad \Delta^l\lambda_\nu=o(\nu^{-l}),\\ \text{(iii)}\quad \sum_{\nu=0}^{\infty}(\nu+1)^l\Delta^{l+1}\lambda_\nu<\infty, & \text{(iv)}\quad \sum_{\nu=0}^{n}(\nu+1)^{l+s}\Delta^{l+1}\lambda_\nu=o(n^s),\\ & \qquad\qquad s>0,\end{array}$$

for any $l=0,1,2,\cdots,k$. Since $\Delta^{k+1}\lambda_\nu\geq 0$, that is $\Delta^k\lambda_\nu\geq\Delta^k\lambda_{\nu+1}$, the sequence $\{\Delta^k\lambda_\nu\}$ decreases to zero. Hence $\Delta^k\lambda_\nu\geq 0$. This implies $\Delta^{k-1}\lambda_\nu\geq 0$ and so on. Finally we get $\Delta\lambda_\nu\geq 0$. Since $\lambda_\nu\to 0$, $\Delta\lambda_\nu\geq 0$, we have

$$(4.4)\qquad \sum_{\nu=0}^{\infty}\Delta\lambda_\nu=\lambda_0,$$

from which, on account of the fact $\Delta^2\lambda_\nu\geq 0$ (that is $\{\Delta\lambda_\nu\}$ decreases to zero), we conclude that $\nu\Delta\lambda_\nu=o(1)$. Summation by parts gives

$$\sum_{\nu=0}^{n}\Delta\lambda_\nu=\sum_{\nu=0}^{n}(A_\nu^{(1)}-A_{\nu-1}^{(1)})\Delta\lambda_\nu=\sum_{\nu=0}^{n-1}A_\nu^{(1)}\Delta^2\lambda_\nu+A_n^{(1)}\Delta\lambda_n,$$

which gives

$$(4.5)\qquad \sum_{\nu=0}^{\infty}A_\nu^{(1)}\Delta^2\lambda_\nu=\lambda_0,$$

since $A_n^{(1)}\Delta\lambda_\nu = O(n\Delta\lambda_n) = 0(1)$. In general, as a consequence of

$$\sum_{\nu=0}^{\infty} A_\nu^{(l-1)}\Delta^l\lambda_\nu = \lambda_0, \qquad 1 \leq l < k+1, \tag{4.6}$$

together with the fact $\Delta^{l+1}\lambda_\nu \geq 0$ we have

$$1/n \sum_{\nu=0}^{n} \nu A_\nu^{(l-1)}\Delta^l\lambda_\nu \to 0,^{3}$$

$$\begin{aligned}
1/n \sum_{\nu=0}^{n} \nu A_\nu^{(l-1)}\Delta^l\lambda_\nu &\geq \Delta^l\lambda_n/n \sum_{\nu=0}^{n} \nu A_\nu^{(l-1)} = \Delta^l\lambda_n/n \sum_{\nu=0}^{n} \nu(A_\nu^{(l)} - A_{\nu-1}^{(l)}) \\
&= \Delta^l\lambda_n/n(-\sum_{\nu=0}^{n-1} A_\nu^{(l)} + nA_n^{(l)}) = (1 - A_{n-1}^{(l+1)}/nA^{(l)})A_n^{(l)}\Delta^l\lambda_n \\
&= l/(l+1)^{-1}A_n^{(l)}\Delta^l\lambda_n,
\end{aligned}$$

whence we conclude that $n^l\Delta^l\lambda_n = o(1)$. Summation by parts gives

$$\sum_{\nu=0}^{n} A_\nu^{(l-1)}\Delta^l\lambda_\nu = \sum_{\nu=0}^{n} (A_\nu^{(l)} - A_{\nu-1}^{(l)})\Delta^l\lambda_\nu = \sum_{\nu=0}^{n-1} A_\nu^{(l)}\Delta^{l+1}\lambda_\nu + A_n^{(l)}\Delta^l\lambda_n,$$

which gives

$$\sum_{\nu=0}^{\infty} A_\nu^{(l)}\Delta^{l+1}\lambda_\nu = \lambda_0, \tag{4.7}$$

so that (ii) and (iii) of (4.3) hold in general for $l = 0, 1, 2, \cdots, k$.

Finally, as is easily seen, since the terms in the (iii) of (4.3) are positive,

$$\sum_{\nu=0}^{n} (\nu+1)^{l+s}\Delta^{l+1}\lambda_\nu = o(n^s) \tag{4.8}$$

for every $s > 0$. Thus (iv) of (4.3) is established.

5. With reference to (4.2), we shall prove in this section that $W_{n\,2}(\theta) = o(1)$, $W_{n\,3}(\theta) = o(1)$ uniformly in θ, and $W_{n\,1}(\theta)$ tends uniformly to a limit. Noticing the definitions of $W_{n\,2}(\theta)$ and $W_{n\,3}(\theta)$ in (4.2), using the fact (iv) of (4.3), we get

$$\begin{aligned}
W_{n\,2}(\theta) &= O(n^{-k}\sum_{p=1}^{k}\sum_{q=1}^{p}\sum_{\nu=0}^{n-(k+1)} (n-\nu)^{k-q}(\nu+1)^{k-p+q}\Delta^{k+1-p}\lambda_{\nu+p}) \\
&= O(\sum_{p=1}^{k}\sum_{q=1}^{p} n^{-q}\sum_{\nu=0}^{n-(k+1)} (\nu+1)^{k-p+q}\Delta^{k+1-p}\lambda_{\nu+p}) = o(1),
\end{aligned} \tag{5.1}$$

$$\begin{aligned}
W_{n\,3}(\theta) &= O(n^{-k}\sum_{q=1}^{k}\sum_{\nu=0}^{n-(k+1)} (n-\nu)^{k-q}(\nu+1)^{q-1}\lambda_{\nu+k+1}) \\
&= O(\sum_{q=1}^{k} n^{-q}\sum_{\nu=0}^{n-(k+1)} (\nu+1)^{q-1}\lambda_{\nu+k+1}) = o(1),
\end{aligned} \tag{5.2}$$

---

[3] If $\sum a_\nu$ converges, then $1/n \sum_{\nu=0}^{n} \nu a_\nu \to 0$.

since $q > 0$. Also

$$(5.3)\quad W_{n\,1}(\theta) = 1/A_n^{(k)} \sum_{p=0}^{k} C_{k+1\,p} C_{p\,0} \sum_{\nu=0}^{n-(k+1)} A_{n-\nu}^{(k)} e^{i\nu\theta} \Delta^p \nu^k \Delta^{k-p+1} \lambda_{\nu+p}$$

$$= 1/A_n^{(k)} \sum_{p=0}^{k} C_{k+1\,p} C_{p\,0} \sum_{\nu=0}^{n} A_{n-\nu}^{(k)} e^{i\nu\theta} \Delta^p \nu^k \Delta^{k-p+1} \lambda_{\nu+p} + o(1),$$

since we have added only a finite number of terms

$$1/A_n^{(k)} \sum_{p=0}^{k} C_{k+1\,p} C_{p\,0} \sum_{\nu=n-(k+1)}^{n} A_{n-\nu}^{(k)} e^{i\nu\theta} \Delta^p \nu^k \Delta^{k-p+1} \lambda_{\nu+p},$$

the order of each term is of the form $O(\nu^{k-p}\Delta^{k-p+1}\lambda_{\nu+p})$, $n-(k+1) \leq \nu \leq n$, $p = 0, 1, \cdots, k$, which is $o(1)$ for $p = 0$ (using (iii) of (4.3) with $l = k$, since $\sum_{\nu=0}^{\infty} (\nu+1)^k \Delta^{k+1}\lambda_\nu < \infty$, so that $\nu^k\Delta^{k+1}\lambda_\nu = o(\mathrm{v})$), and is $o(1/n)$ for $p = 1, \cdots, k$ (by (ii) of (4.3)).

We see that (5.3) consists of $k+1$ terms of the form

$$1/A_n^{(k)} \sum_{\nu=0}^{n} A_{n-\nu}^{(k)} e^{i\nu\theta} \Delta^p \nu^k \Delta^{k-p+1} \lambda_{\nu+p}, \qquad k \geq p,$$

which are the k-th Cesàro means of the absolutely and uniformly convergent series

$$\sum_{\nu=0}^{n} e^{i\nu\theta} \Delta^p \nu^k \Delta^{k-p+1} \lambda_{\nu+p},$$

since $\sum_{j=0}^{\infty} (\nu+1)^{k-p} \Delta^{k-p+1} \lambda_{\nu+p} < +\infty$, by virtue of (iii) of (4.3). Therefore (5.3) is uniformly convergent. (5.1), (5.2) and (5.3) together with (4.1) imply (4.2). Our theorem follows from (2.1), (3.4), (3.5) and (4.1).

The conditions $\lambda_\nu \to 0$ and $\Delta\lambda_\nu \geq 0$ do not imply the uniform $(C, 1)$ summability of the first derived series of (1.1) in the interval $(\epsilon, \pi - \epsilon)$.[4] In the same way we can show that the conditions $\lambda_\nu \to 0$ and $\Delta^k\lambda_\nu \geq 0$ do not imply the uniform (C, k) summability of the k-th derived series of (1.1).

I wish to express my gratitude to Professor Zygmund for his suggestions and encouragement.

UNIVERSITY OF CALIFORNIA.

[4] A. Zygmund, *op. cit.*, p. 129, Ex. 6.

ON ISOLATED EIGENFUNCTIONS ASSOCIATED WITH BOUNDED POTENTIALS.*

By C. R. PUTNAM.

1. Let $f(t)$ be a real-valued, continuous function on the half-line $0 \leqq t < \infty$ and let λ denote a real parameter, $-\infty < \lambda < \infty$. Only real-valued solutions $x = x(t) \not\equiv 0$ of the differential equation

$$x'' + (\lambda + f(t))x = 0 \tag{1}$$

will be considered. If, for some λ, the equation (1) possesses at least one solution $x = x(t)$ not of class (L^2), that is, a solution which *fails* to satisfy

$$\int_0^\infty x^2(t)\,dt < \infty, \tag{2}$$

then, for *every* λ, the equation possesses at least one such solution; [7], p. 238. In this case, (1) is said to be in the *Grenzpunktfall* and the equation (1) and a homogeneous boundary condition

$$x(0)\cos\alpha + x'(0)\sin\alpha = 0, \qquad 0 \leqq \alpha < \pi, \tag{3}$$

determine a boundary value problem for every fixed α. By $S = S(\alpha)$ will be meant the (closed) set of λ-values constituting the spectrum of such a boundary value problem. The derivative of $S(\alpha)$, that is, the set of cluster points of $S(\alpha)$ is independent of α ([7], p. 251), and will be denoted by S'.

It is known ([7], p. 238) that if $f(t)$ is bounded, that is, if

$$|f(t)| < \text{const.}, \qquad 0 \leqq t < \infty, \tag{4}$$

or, more generally, if $f(t)$ is subject only to the unilateral restriction

$$-\infty < f(t) < \text{const.}, \qquad 0 \leqq t < \infty, \tag{5}$$

then (1) is in the Grenzpunktfall. If $f(t)$ satisfies the limit relation

$$f(t) \to 0, \qquad t \to \infty, \tag{6}$$

then the set S' is the half-line $\lambda \geqq 0$; [2], p. 71. In fact, if f is subject only to (4), it follows from [3], p. 850, that every value λ in S' satisfies the inequality $\lambda \geqq -\limsup f(t)$, where $t \to \infty$; furthermore, every λ-interval

* Received April 4, 1949.

$\mu_1 \leqq \lambda \leqq \mu_2$ of length not less than $\limsup f(t) - \liminf f(t)$ and for which μ_1 satisfies $\mu_1 \geqq -\limsup f(t)$, contains at least one point of S' ([5], p. 613).

If (4) is satisfied, it follows from the theorem in [9], p. 6 (cf. also [1]), that if $x(t)$ is a solution of (1) belonging to class (L^2), then $x'(t)$ also is of class (L^2). In **2**, **5**, and **6**, the following criterion for points of $S(\alpha)$ and S', in terms of the solutions of the differential equation (1), will be proved:

THEOREM (I). *Let $f(t)$ be a continuous function on the half-line $0 \leqq t < \infty$ satisfying (4); let λ denote a fixed number for which either of the inequalities*

$$(7) \qquad \lambda + \liminf_{t\to\infty} f(t) > 0; \qquad\qquad (7\text{ bis}) \quad \lambda + \limsup_{t\to\infty} f(t) < 0$$

is satisfied; finally, let $x = x(t) \not\equiv 0$ denote any solution of (1) satisfying (3) for a fixed α. (i) *If*

$$(8) \qquad \limsup_{t\to\infty} \int_0^t (x^2(s) + x'^2(s))ds/(x^2(t) + x'^2(t)) = \infty,$$

then λ is in the set $S(\alpha)$. (ii) *If $x(t)$ is of class (L^2) and if*

$$(9) \qquad \limsup_{t\to\infty} \int_t^\infty (x^2(s) + x'^2(s))ds/(x^2(t) + x'^2(t)) = \infty,$$

then λ is in S'.

In the proof of Theorem (I), it will be convenient to replace assumption (7) by the (apparently more restrictive but, actually, equivalent) assumption

$$(10) \qquad \lambda > \limsup_{t\to\infty} |f(t)|.$$

That (7) and (10) are equivalent follows from the fact that the differential equation (1) remains unchanged if λ and $f(t)$ are replaced by $\lambda + c$ and $f(t) - c$ ($c = \text{const.}$) respectively. Hence (cf. (4)), there is no loss of generality in supposing that

$$-\liminf_{t\to\infty} f(t) = \limsup_{t\to\infty} f(t) \qquad (= \limsup_{t\to\infty} |f(t)|);$$

and, consequently, (7) becomes identical with (10).

If (1) is in the Grenzpunktfall and if λ is not in the set S', then there exists one and (except for constant multiples) only one solution $x = y(t)$ of (1) belonging to class (L^2); [4].

As a partial corollary of Theorem (I), there will be proved

THEOREM (II). *Let $f(t)$ be a continuous function on the half-line $0 \leqq t < \infty$ satisfying* (4); *let λ denote a fixed number not in the set S' and satisfying* (7) *or* (7 bis); *finally, let $x = y(t)$ be a solution of* (1) *belonging to class* (L^2). *Then there exist two positive constants, v and k, satisfying*

$$y^2(t) + y'^2(t) < ve^{-kt}, \qquad 0 \leqq t < \infty. \tag{11}$$

Furthermore, if $x = x(t)$ denotes any solution of (1) *which is not a constant multiple of $y(t)$, then*

$$x^2(t) + x'^2(t) > we^{kt}, \qquad 0 \leqq t < \infty, \tag{12}$$

where w denotes a positive constant.

If $p(t)$ denotes a continuous periodic function on $0 \leqq t < \infty$, it is known ([10] and [6], p. 844) that the set S' associated with the differential equation

$$x'' + (\lambda + p(t))x = 0 \tag{13}$$

is identical with the region of stability of the same equation. In case $f(t)$ satisfies (4) while λ satisfies (7) or (7 bis) and is not in S', it is seen from Theorem (II) that the solutions of (1) behave as the solutions of (13) in the regions of instability; that is, in both cases, there exist (exponentially) "large" and "small" solutions.

It is known ([11], p. 604) that if $f(t)$ satisfies (5) and if λ is not in S', then the "isolated" eigenfunction $y(t)$ belonging to λ (cf. the remark preceding the statement of Theorem (II)) satisfies $y(t) = O(t^{-n})$, $t \to \infty$, for every positive constant n. According to Theorem (II), the last relation can be sharpened to the exponential estimate of (11) provided assumption (5) is strengthened to (4) and the additional assumption, either (7) or (7 bis), is made. It remains undecided, however, whether these altered hypotheses are necessary for this improved estimate.

If (4) holds and λ is arbitrary, it is known ([8], p. 391) that *any* solution $x(t)$ of (1) satisfies, for large t, the inequalities

$$e^{-k_1 t} < x^2(t) + x'^2(t) < e^{k_2 t},$$

for some pair of positive constants k_1 and k_2. It follows from a remark of Wintner [11], p. 604, that the more precise formulation of the above inequalities, given in [8], p. 391, together with (11), implies the theorem of [2] concerning S' in the case (6).

2. *Proof of Theorem* (I) *under assumption* (7 bis). Let $h(t)$

$= -(\lambda + f(t))$ and choose, in virtue of (4) and (7 bis), a constant b such that

(14) $\qquad h(t) > b > 0$, when t is sufficiently large.

Relation (7 bis) implies that (1) is non-oscillatory and consequently λ is not in S'; cf., e. g., [3]. If, therefore, $x = y(t)$ is the (essentially unique) solution of (1) belonging to class (L^2), it follows from [9] (or directly, cf. also [1]), that

$$y(t) \to 0,\ y'(t) \to 0, \qquad t \to \infty. \tag{15}$$

The identity $(yy')' = y'^2 + yy''$ and (1) yield $(yy')' = y'^2 + hy^2$; hence, by (14), if t is sufficiently large, $y'^2 + by^2 < (yy')'$. An integration of this inequality and an application of (15) now imply

$$\int_t^\infty (y'^2 + by^2)\,ds < -y(t)y'(t),$$

if t is sufficiently large. This inequality and the inequality $|\,yy'\,| \leqq \frac{1}{2}(y^2 + y'^2)$ clearly imply

$$\limsup_{t\to\infty} \int_t^\infty (y^2 + y'^2)\,ds/(y^2(t) + y'^2(t)) < \infty.$$

This verifies the fact that the assumption (9) of (ii) is never satisfied in the case (7 bis).

On the other hand, it follows from (15) that (8) holds if $x = y(t)$ is of class (L^2). Let $x = x(t)$ be any solution of (1) linearly independent of $y(t)$. The Wronskian $x'y - xy'$ is a non-vanishing constant; consequently,

$$0 < \text{const.} = |\,x'y - xy'\,| \leqq (x^2 + x'^2)(y^2 + y'^2). \tag{16}$$

Hence, by (15), it is seen that $x^2 + x'^2 \to \infty$ as $t \to \infty$. As above, it is easily verified that, for large t-values,

$$\text{const.} + \int_0^t (x'^2 + bx^2)\,ds < |\,x(t)x'(t)\,|;$$

where the "const." represents a contribution of two sources, namely one related to the lower limit of integration, the other related to the fact that the inequality in (14) is assumed to be valid for large t-values. Since $|\,xx'\,| \leqq \frac{1}{2}(x^2 + x'^2) \to \infty$ as $t \to \infty$, the limit relation (8) is violated. Consequently, (8) holds if and only if $x = x(t)$ is of class (L^2). The assertion (i) is contained in this statement; and so, Theorem (I) is proved in the case (7 bis).

3. Before beginning the proofs of the remainder of Theorem (I) and of Theorem (II), it will be convenient to obtain an inequality from the "Parseval" identity for the boundary value problem determined by (1) and (3). Let $\phi(t,\lambda)=\phi(t,\lambda,\alpha)$ denote the solution of the differential equation (1) satisfying the boundary condition (3) and normalized by

$$\phi(0,\lambda)=\sin\alpha, \qquad \phi'(0,\lambda)=-\cos\alpha,$$

where the prime denotes partial differentiation with respect to t. If $\lambda=\lambda_j$ is an eigenvalue, the symbol $\phi_j(t)=\text{const.}\,\phi(t,\lambda_j)$ will denote an eigenfunction of λ_j normalized so that the integral of $\phi_j{}^2(s)$ over $0\leqq s<\infty$ is 1. If $\rho=\rho(\lambda)$ denotes the unique continuous function normalized by $\rho(0)=0$, determining the continuous spectrum of the boundary value problem (1) and (3), the eigendifferentials $dP(t,\lambda)=dP(t,\lambda,\alpha)$ are defined by

$$P(t,\lambda)=\int_0^\lambda \phi(t,\mu)\,d\rho(\mu), \text{ i. e., } dP(t,\lambda)=\phi(t,\lambda)\,d\rho(\lambda).$$

An eigenfunction ϕ_j satisfies the differential equation

$$\phi''_j+(\lambda_j+f(t))\phi_j=0, \tag{17}$$

while the eigendifferentials $dP(t,\lambda)$ satisfy $(dP)''+(\lambda+f(t))dP=0$, that is,

$$(\Delta P)''+f(t)\Delta P+\int_\lambda^{\lambda+\Delta\lambda}\mu d_\mu P(t,\mu)=0. \tag{18}$$

If $x(t)$ denotes any function of class (L^2) on $0\leqq t<\infty$, the Fourier "coefficients" c_j and $\Delta\Gamma(\lambda)$ are defined by

$$c_j=\int_0^\infty x(s)\phi_j(s)\,ds, \qquad \Delta\Gamma(\lambda)=\int_0^\infty x(s)\Delta P(s,\lambda,\alpha)\,ds. \tag{19}$$

The set of eigenfunctions and eigendifferentials forms a complete orthonormal system on $0\leqq t<\infty$; thus, the Parseval relation

$$\int_0^\infty x^2(s)\,ds=\sum_j c_j{}^2+\int_{-\infty}^\infty (d\Gamma)^2/d\rho \tag{20}$$

is valid.

Let $f(t)$ and $F(t)$ denote continuous functions on $0\leqq t<\infty$ satisfying (4) and

$$|F(t)|<\text{const.}, \qquad 0\leqq t<\infty, \tag{21}$$

respectively. Let $g(t)$ be a continuous function of class (L^2) on $0\leqq t<\infty$. Finally, let $x(t)$ denote a solution of the equation

(22) $$x'' + (\lambda + F(t))x = g(t)$$

belonging to class (L^2) and satisfying (3) for a fixed α, provided that such a solution exists. Consider the boundary value problem determined by (1) and (3) for this value α. On multiplying (22) by ϕ_j and (17) by x, subtracting the resulting equations and then integrating, it is seen that

$$(x'\phi_j - x\phi'_j)\Big|_0^t + \int_0^t (\lambda - \lambda_j + F - f)x\phi_j ds = \int_0^t g\phi_j ds.$$

Since $x(t)$ and $g(t)$ are of class (L^2) and (4), (21), hold, it is clear from (22) that the function $x'' + f(t)x$ is of class (L^2). It follows from a remark of Weyl ([7], pp. 241-242; cf. also [9] and [1]) that

$$x(t)\phi_j'(t) - x'(t)\phi_j(t) \to 0, \qquad t \to \infty.$$

Since $x(t)$ satisfies (3) it follows from (19) and the last two formula lines that

$$\int_0^\infty [(f - F)x + g]\phi_j ds = (\lambda - \lambda_j)c_j.$$

A similar calculation in which (17) is replaced by (18) shows that

$$\int_0^\infty [(f - F)x + g]\Delta P\, ds = \int_\lambda^{\lambda+\Delta\lambda} (\lambda - \mu)\, d\Gamma(\mu),$$

where $\Delta\Gamma$ is defined by (19) (cf. [5], p. 616). The last two formula lines and the Parseval relation (20) applied to the function $(f - F)x + g$ then yield

$$\int_0^\infty [(f - F)x + g]^2 ds = \sum_j (\lambda - \lambda_j)^2 c_j^2 + \int_{-\infty}^\infty (\lambda - \mu)^2 (d\Gamma)^2 d\rho(\mu).$$

In virtue of (20) and the inequality $(a + b)^2 \leqq 2(a^2 + b^2)$, it follows from the last relation that

(23) $$2\left(\int_0^\infty (F - f)^2 x^2 ds + \int_0^\infty g^2 ds\right) \geqq m^2 \int_0^\infty x^2 ds,$$

where $m = m(\lambda, \alpha)$ is defined by

(24) $$m = \min|\lambda - \mu|, \qquad \mu \text{ in } S(\alpha).$$

4. It will be shown that

(*) *If $T > 0$, there exist three positive constants d, c_1 and c_2 (all independent of T); a continuous function $g(t)$ on $0 \leqq t < \infty$ satisfying*

(25) $$g(t) = 0 \text{ for } 0 \leqq t \leqq T \text{ and } T + d \leqq t < \infty;$$

and a function $\psi(t)$ satisfying the differential equation

(26) $$\psi'' + \lambda\psi = g(t),$$

the relations

(27) $$\psi(T) \neq 0, \qquad \psi'(T) = 0, \qquad \psi(T+d) = \psi'(T+d) = 0$$

and, finally,

(28) $$\int_T^{T+d} \psi^2(s)\,ds = c_1\psi^2(T), \qquad \int_0^\infty g^2(t)\,dt = c_2\psi^2(T).$$

Let T_1, T_2 denote a pair of numbers satisfying

(29) $$\pi\lambda^{-\frac{1}{2}} < T_1 < T_2, \qquad T_2 - T_1 < \tfrac{1}{2}\pi\lambda^{-\frac{1}{2}}$$

and let $G(t)$ denote any continuous function on $0 \leqq t < \infty$ satisfying

(30) $G(t)$ is or is not 0 according as $T_1 < t < T_2$ is not or is satisfied.

If the function $\Psi(t)$ is defined by

$$\Psi(t) = \lambda^{-\frac{1}{2}} \int_{T_2}^t G(s) \sin \lambda^{\frac{1}{2}}(t-s)\,ds, \qquad 0 \leqq t < \infty,$$

it is seen that

(31) $$\Psi'(t) = \int_{T_2}^t G(s) \cos \lambda^{\frac{1}{2}}(t-s)\,ds,$$

and, consequently,

(32) $$\Psi(T_2) = \Psi'(T_2) = 0.$$

It is easily verified that $\Psi(t)$ satisfies the differential equation

(33) $$\Psi'' + \lambda\Psi = G(t).$$

In virtue of (29) and (30) the integrand of (31) does not vanish for $T_1 < t < T_2$ and hence $\Psi'(t) \neq 0$ when $T_1 \leqq t < T_2$. It follows from (30) that on the domain $0 \leqq t \leqq T_1$ the solution $\Psi(t)$ of (33) is a non-trivial linear combination of $\sin \lambda^{\frac{1}{2}}t$ and $\cos \lambda^{\frac{1}{2}}t$. Consequently, there exists a (unique) point T_3 such that

(32 bis) $$\Psi(T_3) \neq 0, \qquad \Psi'(T_3) = 0, \qquad T_1 - \pi\lambda^{-\frac{1}{2}} < T_3 < T_1.$$

Define the positive constants c_1 and c_2 by

(34) $$c_1 = \int_{T_3}^{T_2} \Psi^2(s)\,ds/\Psi^2(T_3) \text{ and } c_2 = \int_0^\infty G^2(s)\,ds/\Psi^2(T_3),$$

and d by $T_2 - T_3$.

If $T > 0$ is arbitrary, define the functions $g(t)$ and $\psi(t)$ on $0 \leqq t < \infty$ by

$$g(t) = 0 \text{ for } 0 \leqq t \leqq T \text{ and } g(t) = G(t - T + T_3) \text{ for } T \leqq t < \infty,$$

and

$$\psi(t) = \lambda^{-\frac{1}{2}} \int_{T+d}^{t} g(s) \sin \lambda^{\frac{1}{2}}(t - s)\, ds, \qquad 0 \leqq t < \infty;$$

so that T now plays the rôle of T_3 in (34) and the formula preceding it. Relations (25), (26), (27) and (28) follow from (30), (33), (32), (32 bis) and (34), in that order, and the proof of (*) is complete.

5. *Proof of* (i) *of Theorem* (I) *under assumption* (7). According to the remark following the statement of Theorem (I), it may be assumed that (10) holds. It will be shown that there exists a sequence $t_1 < t_2 < \cdots$, where $t_n \to \infty$ as $n \to \infty$, such that

$$x'(t_n) = 0, \qquad n = 1, 2, \cdots, \tag{35}$$

and

$$\int_0^{t_n} x^2(s)\, ds / x^2(t_n) \to \infty, \qquad n \to \infty. \tag{36}$$

In virtue of (10), a pair of constants β and S can be chosen so that

$$|f(t)| < \beta < \lambda, \qquad t \geqq S. \tag{37}$$

Since $\lambda + f(t) > 0$, when $t \geqq S$, the graph of $x = |x(t)|$ on this domain consists of a sequence of convex arches. If $\tau_1 < \tau_2$, where $\tau_1 \geqq S$, denote two successive zeros of $x'(t)$ it is clear from (37) that

$$\tau_2 - \tau_1 \leqq 2\pi(\lambda - \beta)^{-\frac{1}{2}}. \tag{38}$$

If $X(t) = x^2(t) + x'^2(t)$, the proof of the inequality of [8], p. 391, together with (38) and (4), shows that

$$|\log X(u_1)/X(u_2)| \leqq \gamma, \qquad u_1 \text{ and } u_2 \text{ arbitrary in } [\tau_1, \tau_2],$$

where γ is a constant (depending on λ) independent of the choice of $\tau_1 (\geqq S)$. In particular, the last formula line implies

$$x^2(\tau_2) \leqq e^{\gamma}(x^2(t) + x'^2(t)), \qquad \tau_1 \leqq t \leqq \tau_2,$$

and therefore,

$$\int_0^t X(s)\, ds / X(t) \leqq e^{\gamma} \int_0^{T_2} X(s)\, ds / x^2(\tau_2), \qquad \tau_1 \leqq t \leqq \tau_2.$$

The last inequality and (8) clearly imply the existence of a sequence $t_1 < t_2 < \cdots$, where $t_n \to \infty$, as $n \to \infty$, such that (35) and

$$\int_0^{t_n} (x^2 + x'^2)\, ds / x^2(t_n) \to \infty, \qquad n \to \infty,$$

hold. Multiplication of (1) by x followed by an integration and an application of (35) shows that for $n = 1, 2, \cdots$,

$$\int_0^{t_n} x'^2(s)\, ds = -x(0)x'(0) + \int_0^{t_n} (\lambda + f) x^2 ds. \tag{39}$$

In virtue of (4), the last two relations imply $(A + B \int_0^{t_n} x^2(s) ds)/x^2(t_n) \to \infty$, $n \to \infty$, where A and B are positive constants. This last relation obviously implies (36).

Let D denote an arbitrary positive constant and choose a number N, depending on D, such that (cf. (37))

$$t_{N-1} > S \tag{40}$$

and (cf. (36))

$$\int_0^{t_n} x^2(s)\, ds / x^2(t_N) > D. \tag{41}$$

Let s_N denote the first zero of $x(t)$ to the left of t_N and choose u_N, where $s_N < u_N < t_N$, so near t_N that

$$| x'(u_N) | \leqq | x(u_N) | (\lambda - \beta)^{\frac{1}{2}} / \pi, \tag{42}$$

where β is defined by (37), and

$$\int_0^{u_N} x^2(s)\, ds / x^2(u_N) > D/2. \tag{43}$$

Let $\xi(t)$ denote the solution of the differential equation

$$\xi'' + (\lambda + \beta)\xi = 0 \tag{44}$$

which satisfies

$$\xi(u_N) = x(u_N), \qquad \xi'(u_N) = x'(u_N). \tag{45}$$

Let R $(> u_N)$ denote the first zero of the function $\xi'(t)$ to the right of u_N and define a continuous function $F = F(t)$ on $0 \leqq t < \infty$ so as to satisfy

$$\begin{gathered} F(t) = f(t) \text{ for } 0 \leqq t \leqq u_N, \qquad | F(t) | \leqq | f(t) | \text{ for } u_N \leqq t \leqq R, \\ F(t) = 0 \text{ for } R \leqq t < \infty. \end{gathered} \tag{46}$$

Let $y = y(t)$ denote the solution of the differential equation

$$y'' + (\lambda + F(t))y = 0 \tag{47}$$

which satisfies

$$y(u_N) = x(u_N), \qquad y'(u_N) = x'(u_N). \tag{48}$$

In virtue of (37), (40) and (46) it follows that

$$\lambda - \beta < \lambda + F(t) < \lambda + \beta, \qquad u_N \leqq t < \infty. \tag{49}$$

Hence, if T denotes the first zero of the function $y'(t)$ to the right of u_N, relations (44) to (49) imply the inequalities

$$u_N < R < T. \tag{50}$$

It is easily verified, as a consequence of (48) and (49) that

$$|\, y(T) \,| \leqq |\, x(u_N) \,| + |\, x'(u_N) \,| \, (T - u_N) \text{ and } T - u_N \leqq \pi(\lambda - \beta)^{-\frac{1}{2}}.$$

The last two relations and (42) imply $|\, y(T) \,| \leqq 2 \,|\, x(u_N) \,|$. Since (46), (47) and (48) show that $y(t) \equiv x(t)$ for $0 \leqq t \leqq u_N$, it follows from (43) and (50) that

$$\int_0^T y^2(s)\, ds / y^2(T) > D/8. \tag{51}$$

Identify the point T, just constructed, with that occuring in the italicized statement (*) of this section. Since the assertions of (*) remain unchanged if $g(t)$ and $\psi(t)$ are replaced by $Cg(t)$ and $C\psi(t)$ respectively, where C is any non-vanishing constant, it may be supposed that the function $\psi(t)$ satisfies

$$\psi(T) = y(T). \tag{52}$$

In virtue of (27) and the definition of T (in the last paragraph) it follows that

$$y'(T) = \psi'(T) = 0.$$

Define the function $z = z(t)$ on $0 \leqq t < \infty$ by

$$z(t) = y(t) \text{ for } 0 \leqq t \leqq T \text{ and } z(t) = \psi(t) \text{ for } T \leqq t < \infty.$$

The statement (*) together with (46), (49) and the properties of $y(t)$ show that $z(t)$ satisfies the differential equation

$$z'' + (\lambda + F(t))z = g(t) \tag{53}$$

and boundary condition (3), for the (fixed) α specified in the statement of Theorem (I). The third relation of (27) and the facts that $g(t)=0$ when $T+d\leqq t<\infty$ and $F(t)=0$ when $t\geqq R$, where $R<T<T+d$, imply that

$$(54)\qquad z(t)\equiv 0, \qquad T+d\leqq t<\infty.$$

Let the functions $f(t)$, $F(t)$ and $g(t)$, considered above, be identified with those appearing in the latter part of **3**. It follows from (23), (46), (53) and (54) that

$$(55)\qquad 2\Big(\int_{u_N}^{T+d}(F-f)^2z^2ds+\int_0^\infty g^2ds\Big)\geqq m^2\int_0^\infty z^2ds.$$

Hence, by (52) and the second equality of (28), $\int_0^\infty g^2ds=c_2y^2(T)$, and by (4) and (46),

$$(F(t)-f(t))^2<4f^2(t)<c_3\ (=\text{const.}),\qquad 0\leqq t<\infty.$$

Thus $\int_T^{T+d}(F-f)^2z^2dt\leqq c_3c_1y^2(T)$, by the first equality of (28) and the definition of $z(t)$. Since $T-u_N\leqq\pi(\lambda-\beta)^{-\frac{1}{2}}\ (=c_4)$ and $|\,z(t)\,|=|\,y(t)\,|\leqq|\,y(T)\,|$ for $u_N\leqq t\leqq T$, by the definition of u_N and T, it follows that

$$\int_{u_N}^T(F-f)^2z^2\,ds\leqq c_3c_4y^2(T).$$

Relation (55) and the above relations imply

$$c_5y^2(T)\geqq m^2\int_0^\infty z^2ds\geqq m^2\int_0^T y^2ds,$$

where c_5 denotes the constant $c_5=2(c_2+c_1c_3+c_3c_4)$. It follows therefore from (51) that m^2D is less than a constant which is independent of D. Consequently, $m=0$, since D may be chosen arbitrarily large. This means, according to the definition (24) of m, that λ is in $S(\alpha)$ and the proof of (i) is complete.

6. *Proof of* (ii) *of Theorem* (I) *under assumption* (7). By a process similar to that used in **5**, it is easily shown that there exists a sequence $t_1<t_2<\cdots$, where $t_n\to\infty$ as $n\to\infty$, such that (35) and

$$\int_{t_n}^\infty(x^2+x'^2)ds/x^2(t_n)\to\infty,\qquad n\to\infty.$$

In virtue of (4), relations (15) are valid for $x = y$ and consequently, from (35),

$$\int_{t_n}^{\infty} x'^2 ds = \int_{t_n}^{\infty} (\lambda + f) x^2 ds, \qquad n = 1, 2, \cdots,$$

cf. (39). The last two formula lines clearly imply $\int_{t_n}^{\infty} x^2 ds / x^2(t_n) \to \infty$, $n \to \infty$, corresponding to (36). The remainder of the proof of (ii), including the construction of functions corresponding to $g(t)$, $F(t)$, etc. occurring in **4** and **5** is similar to that of (i) provided only that it is observed that the rôles played by 0 and ∞ there may be interchanged in the present case. This procedure is permitted by the assumption that x (and hence x') is of class (L^2), together with the resulting implication (15) for $x = y$. A copying of the proof of (i) given in the last two sections, with the appropriate modifications as indicated, leads to the construction of a function $Z(t)$, corresponding to the function $z(t)$ above, satisfying the identity $Z(t) \equiv 0$, $0 \leqq t \leqq \text{const.} < \infty$, corresponding to (54) for $z(t)$. Since $Z(t)$ satisfies, therefore, the boundary condition (3) for *every* α, the function $m = m(\lambda, \alpha)$ defined by (24) satisfies the identity $m(\lambda, \alpha) \equiv 0$, $0 \leqq \alpha < \pi$. That is, λ is in $S(\alpha)$ for every α and hence, in S'. This completes the proof of (ii) (and hence of Theorem (I)).

7. *Proof of Theorem* (II). It follows from (ii) of Theorem (I) that the expression appearing on the left of the equality (9), in which $x = y$, is finite. Hence, if $Y(t)$ is defined by

$$Y(t) = \int_{t}^{\infty} (y^2 + y'^2) ds, \tag{56}$$

there exists a positive constant k such that $Y(t)/-Y'(t) < 1/k$, $0 \leqq t < \infty$; consequently,

$$Y(t) < Y(0) e^{-kt}. \tag{57}$$

Relations (15) and (1), where $x = y(t)$, imply that

$$y'^2(t) + \lambda y^2(t) = 2 \int_{t}^{\infty} (y'y'' + \lambda y'y) ds = -2 \int_{t}^{\infty} fyy' ds.$$

In virtue of (4), the last relation and the inequality $|yy'| \leqq 2(y^2 + y'^2)$ imply

$$y'^2(t) + \lambda y^2(t) \leqq \text{const.} \int_{t}^{\infty} (y^2 + y'^2) ds, \qquad 0 \leqq t < \infty.$$

The relation (11) follows from this inequality, (56), (57) and $\lambda > 0$.

Relation (12) clearly follows from (11) and (16), and the proof of Theorem (II) is complete. (The existence of a constant $k > 0$ satisfying (12) can also be obtained directly from (i) of Theorem (I).)

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "The L^2-solutions of linear differential equations of second order," *Duke Mathematical Journal*, vol. 14 (1947), pp. 323-326.

[2] ———, "On the spectra of slightly disturbed linear oscillators," *American Journal of Mathematics*, vol. 71 (1949), pp. 71-79.

[3] ——— and C. R. Putnam, "The least cluster point of the spectrum of boundary value problems," *ibid.*, vol. 70 (1948), pp. 849-855.

[4] ——— and A. Wintner, "An oscillation theorem for continuous spectra," *Proceedings of the National Academy of Sciences*, vol. 33 (1947), pp. 376-379.

[5] C. R. Putnam, "The cluster spectra of bounded potentials," *American Journal of Mathematics*, vol. 71 (1949), pp. 612-620.

[6] S. Wallach, "The spectra of periodic potentials," *ibid.*, vol. 70 (1948), pp. 842-848.

[7] H. Weyl, "Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen," *Mathematische Annalen*, vol. 68 (1910), pp. 222-269.

[8] A. Wintner, "The adiabatic linear oscillator," *American Journal of Mathematics*, vol. 68 (1946), pp. 385-397.

[9] ———, "(L^2)-connections between the potential and kinetic energies of linear systems," *ibid.*, vol. 69 (1947), pp. 5-13.

[10] ———, "Stability and spectrum in the wave mechanics of lattices," *Physical Review*, vol. 72 (1947), pp. 81-82.

[11] ———, "On the smallness of isolated eigenfunctions," *American Journal of Mathematics*, vol. 71 (1949), pp. 603-611.

ON THE DERIVATIVES OF THE SOLUTIONS OF ONE-DIMENSIONAL WAVE EQUATIONS.*

By PHILIP HARTMAN and AUREL WINTNER.

1. On the half-line $0 \leqq t < \infty$, consider the linear differential equation

$$x'' + q(t)x = 0, \tag{1}$$

where $q(t)$ is real-valued and continuous. By solutions of (1) will be meant real-valued solutions $x(t) \not\equiv 0$. For such a solution, and its derivative, the (L^2)-character is defined by the respective conditions

$$(2_1) \quad \int_0^\infty x^2(t)\,dt < \infty; \qquad (2_2) \quad \int_0^\infty x'^2(t)\,dt < \infty.$$

The present note centers about the following facts:

For an unspecified $q(t)$ in (1), where $0 \leqq t < \infty$, let $x = x(t)$ and $x = y(t)$ be two linearly independent solutions. Then

(i) *for suitable $q(t)$, both $x(t)$ and $y(t)$ become of class (L^2);*

(ii) *for no $q(t)$ can $x'(t)$ and $y'(t)$ be of class (L^2);*

(iii) *if $x'(t)$ is of class (L^2), then $y(t)$ cannot be of class (L^2).*

Remark. If (1) is generalized to

$$(p(t)x')' + q(t)x = 0, \tag{3}$$

where $p(t)$ is positive and continuous, then (ii) cannot be asserted. For instance, if $p(t) = e^t$ and $q(t) = e^{-t}$, then every solution of (3), being given by $x(t) = c_1 \cos e^{-t} + c_2 \sin e^{-t}$, has a derivative satisfying (2_2). What is true is that, with reference to (3),

$$\int_0^\infty p(t)x'^2(t)\,dt = \infty \text{ holds for some } x(t) \text{ if } \int_0^\infty \{p(t)\}^{-1}dt = \infty. \tag{4}$$

Since (ii) refers to (1), where $p(t) \equiv 1$, it is clear that (ii) is contained

* Received May 21, 1949.

in (4). Conversely, (4) can be obtained from (ii) by a change of the independent variable.

Ad (i). The asymptotic results of [10] imply that, if $q(t)$ is of class C'', tends to ∞ as $t \to \infty$, and satisfies

$$\limsup_{t\to\infty} q'^2/q^3 < \infty \text{ and } \int^{\infty} (q'^2/q^{5/2})\,dt < \infty,$$

then every or no solution of (1) is of class (L^2) according as $\int^{\infty} \{q(t)\}^{-\frac{1}{2}}dt$ is convergent or divergent. On the other hand, both requirements of the last formula line are readily seen to be satisfied whenever $q(t)$ is a logarithmico-exponential function which tends to ∞ as $t \to \infty$. Accordingly, if $q(t)$ is any such function, (i) will hold whenever $q(t)$ increases at least as fast as t^2 (or just $t^2/\log^2 t, \cdots$); cf. [4], p. 306.

Proof of (ii). Since $x(t)$ and $y(t)$ are linearly independent solutions of (1), their Wronskian is a non-vanishing constant. Hence, it can be assumed that

$$xy' - yx' \equiv 1. \tag{5}$$

Then, by Schwarz's inequality, $1 \leqq (x^2 + y^2)(x'^2 + y'^2)$. Consequently, if (ii) is assumed to be false, i. e., if both $x'(t)$ and $y'(t)$ are of class (L^2), it follows that $1/(x^2 + y^2)$ is absolutely integrable (over $0 \leqq t < \infty$). This implies that (1) is non-oscillatory; cf. [5], pp. 210-211. But if (1) is non-oscillatory, then it must possess some solution $x = z(t)$ satisfying

$$\int_T^{\infty} \{z(t)\}^{-2}dt < \infty \text{ if } T \text{ is large enough}; \tag{6}$$

cf. [2], p. 703. On the other hand, since $z(t)$ is a linear combination of $x(t)$ and $y(t)$, and since $x'(t)$ and $y'(t)$ are supposed to be of class (L^2), the function $z'(t)$ is of class (L^2). In view of (6) and of Schwarz's inequality, this implies that the product $z'z^{-1}$ is integrable over the half-line $T \leqq t < \infty$. Since $z'z^{-1} = (\log z)'$, it follows that $\log z(t)$, and therefore $z(t)$ itself, tends to a finite limit as $t \to \infty$. Since this contradicts (6), the proof of (ii) is complete.

Proof of (iii). Suppose that (iii) is false, i. e., that $x'(t)$ and $y(t)$ are of class (L^2). Then, if (5) is written in the form

$$(xy)' - 1 = 2x'y, \tag{7}$$

it is seen that $(xy)'-1$ is (absolutely) integrable over $0 \leqq t < \infty$. Hence, $xy-t$ tends to a finite limit as $t \to \infty$. This implies that, if t is large enough, $x(t)$ and $y(t)$ do not vanish and satisfy $1/|x(t)| < |y(t)|$. Since $y(t)$ is of class (L^2), it follows that $1/x^2(t)$ is absolutely integrable over $T \leqq t < \infty$, if T is large enough (so large that $x(t) \neq 0$ when $t \geqq T$). But (5) shows that $1/x^2(t)$ is identical with the derivative of y/x. Consequently, y/x tends to a finite limit as $t \to \infty$. This limit is 0; for, on the one hand, y is of class (L^2) and, on the other hand, $xy \sim t$ as $t \to \infty$. Hence

$$y(t)/x(t) = -\int_t^\infty ds/x^2(s) \text{ whenever } t \geqq T.$$

Thus, $yx < 0$ if $t \geqq T$. But this contradicts $xy \sim t$, and completes therefore the proof of (iii).

2. A modification of this proof leads to the substantial refinement of (iii). In this connection, the following lemma is of interest:

LEMMA. *If* (1), *where* $0 \leqq t < \infty$, *is non-oscillatory, and if* $x = x(t)$ *and* $y = y(t)$ *are two linearly independent solutions the second of which is of class* (L^2), *then* (6) *is satisfied by* $z = x$.

This Lemma is between the lines of [2], p. 703. It can be verified as follows: Since (1) is non-oscillatory, $x(t) \neq 0$ if t is large enough, say $t \geqq T$. Then $(y/x)' = 1/x^2$, by (5). Hence, if (6) is denied for $z = x$, it follows that $y/x \to \infty$. But this is impossible, since y is of class (L^2) while x is not. For, if x were of class (L^2), then, since y is, (1) had two linearly independent solutions of class (L^2). As shown in [2], this is contradicted by the assumption that (1) is non-oscillatory.

The refinement of (iii), referred to above, is as follows:

(*) *If* (1), *where* $0 \leqq t < \infty$, *has a solution* $x(t)$ *the derivative of which satisfies*

$$\int_0^t x'^2(s)\,ds = O(t^2), \tag{8}$$

then no solution linearly independent of this $x(t)$ *is of class* (L^2) (*while* $x(t)$ *itself may, but need not, be of class* (L^2)).

COROLLARY. *If the derivative of some solution* $x(t) \neq 0$ *of* (1) *satisfies* (8) (*e. g., if*

(8 bis) $\qquad$ either $x'(t) = O(t^{\frac{1}{2}})$ or $\int_0^\infty x'^2(t)\,dt < \infty$

holds for some $x(t) \not\equiv 0$), *then* (1) *cannot have two linearly independent solutions of class* (L^2).

Remark. The $O(t^2)$ in (8) cannot be improved to $O(t^{2+\epsilon})$, where $\epsilon > 0$. In fact, the $O(t^{\frac{1}{2}})$ in (8 bis) cannot be relaxed to $O(t^{\frac{1}{2}+\epsilon})$. In order to see this, it is sufficient to choose $q(t) = t^{2+\epsilon}$ and then apply the general asymptotic results of [10].

Proof of (*). Suppose that (*) is false. Then (1) has a solution $x = y(t)$ which is of class (L^2) and satisfies (5), where $x(t)$ is the solution occurring in (8). Hence, if C and T are large enough,

$$\int_0^t x'^2(s)\,ds < (Ct)^2 \text{ if } T < t < \infty, \text{ and } \int_T^\infty y^2(s)\,ds < (4C)^{-2}.$$

Consequently, Schwarz's inequality shows that, if $T > 1$,

$$2\int_T^t |\, x'(s)y(s)\,|\, ds < \tfrac{1}{2}t \text{ whenever } t > T.$$

Since (7) implies that $x(t)y(t) = \text{const.} + 2\int_T^t x'(s)y(s)\,ds + t$, it follows that

$$x(t)y(t) > \tfrac{1}{2}t + \text{const. if } T < t < \infty. \tag{9}$$

Since $y(t)$ is of class (L^2), it is clear from (9) that $x(t)$ cannot be of class (L^2). It also follows from (9) that $x(t)$, $y(t)$ do not vanish for large t, i.e., that (1) is non-oscillatory. Hence, by the Lemma, (6) is satisfied by $z(t) = x(t)$. Clearly, the balance of the proof of (*) is substantially identical with the end of the above Proof of (iii).

(* bis) *The assertion of* (*) *remains true if* (1) *is generalized to* (3) *but* (8) *is replaced by*

$$\int_0^t x'^2(s)\,ds = O(\int_0^t \{p(s)\}^{-1} ds)^2, \tag{10}$$

provided that the integral on the right of (10) *is not* $O(1)$; cf. (4).

COROLLARY. *If* (3), *where* $0 \leqq t < \infty$, *has a solution the derivative of which satisfies* (10), *and if*

$$\int_0^\infty \{p(s)\}^{-1} ds = \infty, \tag{4 bis}$$

then (3) *cannot have two linearly independent solutions of class* (L^2).

3. Let (1) now be replaced by

$$x'' + \{q(t) + \lambda\}x = 0, \tag{11}$$

where λ is a real parameter. Suppose that (11) is of *Grenzpunkt* type, i. e., that not every solution of (11) is of class (L^2). In order that this be the case, it is sufficient that

$$-\infty \leqq \limsup q(t) < \infty \qquad (t \to \infty) \tag{12}$$

([8], p. 238). Another sufficient condition is the existence of a constant satisfying the unilateral Lipschitz condition

$$q(t_2) - q(t_1) < \text{const.}\,(t_2 - t_1) \text{ for } 0 \leqq t_1 < t_2 < \infty \tag{13}$$

([4], p. 296). The above considerations lead to a more far-reaching result.

In order to describe the situation completely, assign to (11) a linear boundary condition at $t = 0$, e. g., $x(0) = 0$ or, more generally,

$$x(0) \cos \alpha + x'(0) \sin \alpha = 0, \qquad (0 \leqq \alpha < \pi). \tag{14}$$

Denote by $S(\alpha)$ the set of λ-values representing the spectrum which is determined by (11) and (14) when (11) is of *Grenzpunkt* type. According to Weyl [8], p. 251, the set of the cluster values of $S(\alpha)$ is independent of α and can, therefore, be denoted simply by S'. Since $S(\alpha)$ is closed, S' is contained in every $S(\alpha)$. Conversely, a λ is in S' whenever it is in every $S(\alpha)$ (or, for that matter, in $S(\alpha_1)$ and $S(\alpha_2)$, where $\alpha_1 \not\equiv \alpha_2 \bmod \pi$). Let S' be called the essential spectrum of (11).

The following theorem can now be proved:

(I) *Suppose that there exists a* $\lambda = \lambda_0$ *corresponding to which* (11) *has a solution* $x(t) \not\equiv 0$ *satisfying* (8) (*e. g.*, (8 bis)). *Then* (11) *is of Grenzpunkt type. Furthermore, either* $x(t)$ *is of class* (L^2) *or* λ_0 *is in the essential spectrum.*

The second of these two possibilities (which are not mutually exclusive) must occur if (8) *is satisfied by two linearly independent solutions of the case* $\lambda = \lambda_0$ *of* (11).

Proof of (I). According to Weyl [8], p. 238, all solutions of (11) are of class (L^2) for some λ only if the same is true for every λ. Since there is no loss of generality in assuming that $\lambda_0 = 0$, the first assertion of (I) is equivalent to the Corollary of (*).

In order to prove the second assertion of (I), suppose that $\lambda_0 = 0$ is not in the essential spectrum. Then (1) must have a solution of class (L^2); cf. [3]. This solution, say $x = x^*(t)$, is unique to a constant factor, since (1) cannot have two linearly independent solutions of class (L^2). It remains to show that $x^*(t)$ is linearly dependent on the $x(t)$ for which (8) is assumed. But if such were not the case, an application of (*) to $y(t) = x^*(t)$ would lead to a contradiction.

This proves the second assertion of (I). Clearly, the third assertion of (I) follows in the same way as the second. In fact, (*) implies that no (non-trivial) solution of (11) is of class (L^2). Hence, by [3], the value $\lambda = \lambda_0$ is in the essential spectrum.

4. A modification of the preceding arguments leads to the following theorem:

(II) *Suppose that the coefficient function of* (11) *satisfies* (12) *(so that, in particular,* (12) *is of Grenzpunkt type). Suppose further that* (12) *has, for some* $\lambda = \lambda_0$, *a solution* $x = x(t) \not\equiv 0$ *satisfying*

$$(15) \qquad \int_0^t x^2(s)\,ds = O(t^N) \text{ for some } N.$$

Then either $x(t)$ *is of class* (L^2) *or* λ_0 *is in the essential spectrum.*

The second of these two possibilities (which are not mutually exclusive) must occur if (15) *is satisfied by two linearly independent solutions of the case* $\lambda = \lambda_0$ *of* (11).

The assertions of (II) were proved in [6] in the particular case in which $x(t) = O(1)$, rather than just (15), is assumed.

Proof of (II). Without loss of generality, let $\lambda_0 = 0$. Then (1) is

supposed to have a solution $x(t) \not\equiv 0$ satisfying (15). If $x = y(t)$ is any solution of (1) linearly independent of this $x(t)$, then, since the Wronskian $xy' - xy'$ is a non-vanishing constant, the square of this constant will satisfy

$$0 < \text{const.} \leqq (x^2 + x'^2)(y^2 + y'^2) \text{ for } 0 \leqq t < \infty,$$

by Schwarz's inequality. On the other hand, from (15),

$$\int_1^\infty t^{-K} x^2(t)\,dt < \infty$$

holds for some K (in fact, for every $K > N + 1$). But the existence of such a K and the assumption (12) imply that

$$\int_1^\infty t^{-K} x'^2(t)\,dt < \infty$$

holds for the same K; cf. [9], p. 9. In view of the last three formula lines, there exists a $C > 0$ satisfying

$$(16) \qquad \int_1^\infty \{y^2(t) + y'^2(t)\}^{-1} t^{-C} dt < \infty.$$

Suppose that the first assertion of (II) is false, i. e., that $x(t)$ is not of class (L^2) and $\lambda = 0$ is not in the essential spectrum. Then (1) has a solution $x = y(t)$ which is linearly independent of $x(t)$ and is of class (L^2); cf. [3]. But (12) necessitates, for every such $y(t)$, the estimate $y(t) = O(t^{-n})$, where n is arbitrarily large; cf. [11]. It follows therefore by the procedure of [1] that $y'(t) = O(t^{-n})$ holds for every n. This pair of O-estimates contradicts (16), since n can be chosen large enough with reference to a fixed C. This contradiction proves the first assertion of (II).

In the remaining assertion of (II), the assumption is that (15) holds for two linearly independent solutions of (1), say for $x = x(t)$ and $x = y(t)$. Since (16) was deduced from (15), and since $x(t)$ in (15) can now be replaced by $y(t)$, it follows that both (16) and (16′) hold in the present case, if (16′) denotes what results if $y(t)$ in (16) is replaced by $x(t)$. Consequently, (1) has no (non-trivial) solution $x = x(t)$ satisfying the estimates $x(t) = O(t^{-n})$ and $x'(t) = O(t^{-n})$, where n is arbitrarily large. Hence, by

[11], the point $\lambda = 0$ is in the essential spectrum and the proof of (II) is complete.

Remark. If the assumption (12) of (II) is strengthened to

$$|q(t)| < \text{const.} \qquad (0 \leqq t < \infty) \tag{17}$$

and if λ in (11) satisfies

$$|\lambda| > \limsup |q(t)| \qquad (t \to \infty), \tag{18}$$

then (II) remains correct when (15) is replaced by

$$\int_0^t x^2(s)\,ds = O(e^{\epsilon t}) \quad \text{for every } \epsilon > 0. \tag{19}$$

The proof of this remark is similar to that of (II). For, by (19),

$$\int_0^\infty x^2(t)e^{-\epsilon t}dt < \infty$$

for $\epsilon > 0$, and by the arguments of [9], p. 9,

$$\int_0^\infty x'^2(t)e^{-\epsilon t}dt < \infty$$

holds in view of (12). On the other hand, it is shown in [7] that (17) and (18) imply that if $\lambda = \lambda_0$ is not in the essential spectrum, then (11) has a solution $x = y(t)$ satisfying the estimates $y(t) = O(e^{-kt})$ and $y'(t) = O(e^{-kt})$ for some $k > 0$. The proof can now be completed as above.

It remains undecided whether or not the assertion (II) modified by replacing (15) by (19) is true without the additional assumptions (17) and (18).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "The L^2-solutions of linear differential equations of second order," *Duke Mathematical Journal*, vol. 14 (1947), pp. 323-326.

[2] ———, "Differential equations with non-oscillatory eigenfunctions," *ibid.*, vol. 15 (1948), pp. 697-709.

[3] ——— and A. Wintner, "An oscillation theorem for continuous spectrum," *Proceedings of the National Academy of Sciences*, vol. 33 (1947), pp. 376-379.

[4] ——— and A. Wintner, "Criteria of non-degeneracy of the wave equation," *American Journal of Mathematics*, vol. 70 (1948), pp. 295-308.

[5] ——— and A. Wintner, "A criterion for the non-degeneracy of the wave equation," *ibid.*, vol. 71 (1949), pp. 206-213.

[6] ——— and A. Wintner, "On the location of spectra of wave equations," *ibid.*, vol. 71 (1949), pp. 214-217.

[7] C. R. Putnam, "On isolated eigenfunctions associated with bounded potentials," *ibid.*, vol. 72 (1950), pp. 135-147.

[8] H. Weyl, "Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugenhörigen Entwicklungen willkürlicher Funktionen," *Mathematische Annalen*, vol. 68 (1910), pp. 222-269.

[9] A. Wintner, "(L^2) connections between the potential and kinetic energies of linear systems," *American Journal of Mathematics*, vol. 69 (1947), pp. 5-13.

[10] ———, "On the normalization of characteristic differentials in continuous spectra," *Physical Review*, vol. 72 (1947), pp. 516-517.

[11] ———, "On the smallness of isolated eigenfunctions," *American Journal of Mathematics*, vol. 71 (1949), pp. 603-611.

ZUSÄTZLICHE STABILITÄTSBETRACHTUNG BETREFFEND "DIE SYMMETRISCHEN PERIODISCHEN BAHNEN DES RESTRINGIERTEN DREIKÖRPERPROBLEMS IN DER NACHBARSCHAFT EINES KRITISCHEN KEPLERKREISES." *

Von Ernst Hölder.

Einer Anregung von Wintner aus dem Jahr 1938 zufolge habe ich (im Sommer 1945) auch die *Stabilität* der in meiner im Hill-Gedächtnisheft erschienenen Arbeit [1] berechneten periodischen Bahnen untersucht. Dazu ist Λ^* bis zu den Gliedern *vierten Grades* in $\zeta, \zeta', \cdots$ explizit zu berechnen. Man hat jetzt also vollständiger

$$(21') \qquad V_4 = V_3 + 2a^{-1}\zeta^4 = V_3 + a^2 m^{-2} \cdot 2(m+1)^2 \zeta^4.$$

In den beiden ersten Zeilen von (23) erscheint jetzt die Entwicklung des Faktors

$$[1 + 2\zeta + \zeta^2 + \zeta'^2]^{\frac{1}{2}} = (1 + \zeta + \tfrac{1}{2}\zeta'^2 - \tfrac{1}{2}\zeta\zeta'^2 + \tfrac{1}{2}\zeta^2\zeta'^2 - \tfrac{1}{8}\zeta'^4 + \cdots)$$

sowie statt $V_3^{\frac{1}{2}}$

$$(24') \qquad V_4^{\frac{1}{2}} = V_3^{\frac{1}{2}} + am^{-1}\{-\tfrac{1}{8}(m+1)(17m^3 + 15m^2 + 3m - 3)\zeta^4 + \cdots\}.$$

Die Restglieder $\Lambda^{(1)}, \Lambda^{(2)}, \cdots$ enthalten ausser von ζ, ζ' unabhängigen Gliedern solche Glieder dritter Ordnung in κ, μ, ζ, ζ', die den Faktor μ enthalten, Glieder vierter Ordnung, die den Faktor κ enthalten, ferner Glieder fünfter und höherer Ordnung.

In

$$(27) \qquad a^{-2} m^2 \Lambda^* = L$$

braucht man ausser

$$\kappa\Theta^{(2)}(\zeta, \zeta) = \kappa[B_{1020}\zeta^2 + B_{1002}\zeta'^2] \text{ mit } B_{1020} = \tfrac{3}{2} j(3m+1), \; B_{1002} = \tfrac{1}{2}$$

und

$$\Theta^{(3)}(\zeta, \zeta, \zeta) = B_{0030}\zeta^3 + B_{0012}\zeta\zeta'^2 \text{ mit } B_{0030} = -j^2 m,$$
$$B_{0012} = -j = -(m+1)$$

* Received June 2, 1949.

[1] E. Hölder, "Die symmetrischen periodischen Bahnen des restringierten Dreikörperproblems in der Nachbarschaft eines kritischen Keplerkreises," *American Journal of Mathematics*, vol. 60 (1938), pp. 801-814.

noch explizit die weiteren Glieder

$$\Theta^{(4)}(\zeta, \zeta, \zeta, \zeta) = B_{0040}\zeta^4 + B_{0022}\zeta^2\zeta'^2 + B_{0004}\zeta'^2$$

mit

$$B_{0040} = -j^2(17m^2 + 6m + 1)/8, \quad B_{0022} = j(-m + 5)/4, \quad B_{0004} = -1/8.$$

Die Lagrange-Funktion wird dann vollständiger

$$(27') \quad L = a^{-2}m\Lambda^* = \tfrac{1}{2}\zeta'^2 - \tfrac{1}{2}j^2\zeta^2 + \Sigma B_{ijkl}\kappa^i\mu^j\zeta^k\zeta'^l = L^{(2)}(\zeta,\zeta) + \Theta$$
$$= L^{(2)}(\zeta,\zeta) + \kappa\Theta^{(1)}(\zeta) + \kappa^2\bar{\Theta}^{(1)}(\zeta) + \kappa\Theta^{(2)}(\zeta,\zeta) + \Theta^{(3)}(\zeta,\zeta,\zeta)$$
$$+ \Theta^{(4)}(\zeta,\zeta,\zeta,\zeta) + \mu L^{(1)}(\zeta) + \Lambda^{(4)}.$$

Die in ζ, ζ' linearen Glieder $\kappa\Theta^{(1)}(\zeta) + \kappa^2\bar{\Theta}^{(1)}(\zeta) + \mu L^{(1)}(\zeta)$ interessieren uns im folgenden nicht.

Die lineare Variationsgleichung für einen infinitesimalen Zuwachs

$$Z(\tau) = C_1 \cos j\tau + C_2 \sin j\tau + \cdots = C_1 Z_1(\tau) + C_2 Z_2(\tau) + \cdots$$

lautet

$$(L_{\zeta'\zeta'}Z' + L_{\zeta'\zeta}Z)' - (L_{\zeta\zeta'}Z' + L_{\zeta\zeta}Z) = 0.$$

Die an dieser Stelle nicht zu entwickelnde Störungstheorie der charakteristischen Exponenten (analog zu jener der Eigenwerte) liefert einen Instabilitätsbereich zwischen den beiden Kurven in der (κ, ξ)-Ebene, auf denen die Variationsgleichung eine periodische Lösung besitzt. Dafür muss die "rechte Seite" die beiden Orthogonalitätsrelationen bezüglich $Z_i(\tau)$ erfüllen, die für C_1, C_2 zwei lineare homogene Gleichungen

$$-\sum_{j=1}^{2} \int_0^{2\pi} (\Theta_{\zeta'\zeta'}Z'_iZ'_j + \Theta_{\zeta\zeta'}(Z_iZ'_j + Z'_iZ_j) + \Theta_{\zeta\zeta}Z_iZ_j)\,d\tau.\ C_j = 0. \qquad (i = 1, 2)$$

darstellen. Für deren Lösbarkeit ist das Verschwinden der Determinante

$$(*) \qquad \det\Big(\int_0^{2\pi} (\Theta_{\zeta'\zeta'}Z'_iZ'_j + \Theta_{\zeta\zeta'}(Z_iZ'_j + Z'_iZ_j) + \Theta_{\zeta\zeta}Z_iZ_j)\,d\tau \Big) = 0 \qquad (i, j = 1, 2)$$

notwendig und hinreichend.

Wegen der Symmetrie der Ausgangslösung

$$\zeta = 2m^2\kappa/(ja^2) + \xi \cos j\tau + \cdots = \kappa Z_0 + \xi Z_1(\tau) + \cdots$$

zerfällt die Determinantengleichung (*) in die beiden Faktoren $(i = 1, 2)$

$$(**) \qquad \kappa \int_0^{2\pi} (2\Theta^{(2)}(Z_iZ_i) + 3.2\Theta^{(3)}(Z_0Z_iZ_i))\,d\tau$$
$$+ \xi^2.4.3 \int_0^{2\pi} \Theta^{(4)}(Z_1Z_1; Z_iZ_i)\,d\tau + \cdots = 0.$$

Dabei sind die polarisierten Formen in unserem Fall

$$-\Theta^{(3)}(Z_0ZZ) = (m+1)(m(m+1)Z_0Z^2 + \tfrac{1}{3}Z_0Z'^2 + \tfrac{2}{3}Z'_0ZZ')$$

und

$$-\Theta^{(4)}(Z_1Z_1ZZ) = \tfrac{1}{8}(m+1)^2(17m^2+6m+1)Z_1^2Z^2 + \tfrac{1}{8}Z'_1{}^2Z'^2$$
$$-\tfrac{1}{4}(m+1)(-m+5)(\tfrac{1}{6}Z_1^2Z'^2 + \tfrac{1}{6}Z'_1{}^2Z^2 + \tfrac{2}{3}Z_1Z'_1ZZ').$$

Mit Rücksicht auf die Integrale

$$\int_0^{2\pi}\binom{\cos^4}{\sin^4} j\tau d\tau = \tfrac{1}{4}(1+\tfrac{1}{2})2\pi, \qquad \int_0^{2\pi}\cos^2 j\tau \sin^2 j\tau d\tau = \tfrac{1}{4}(1-\tfrac{1}{2})2\pi$$

bekommt man (**) ausgerechnet

$$2\kappa(B_{1020} + j^2B_{1002} + 3B_{0030}Z_0 + j^2B_{0012}Z_0)\pi$$
$$= -4.3\xi^2\{(B_{0040}+j^4B_{0004})(2\pm1) + j^2B_{0022}(\tfrac{2}{6}(2\mp1)\pm\tfrac{2}{3})\}\pi/4$$
$$= -(2\pm1)\xi^2\{3B_{0040} + 3(m+1)^4B_{0004} + (m+1)^2B_{0022}\}\pi.$$

Wegen

$$a^{-2}m^2 = (m+1)^{4/3}m^{2/3} > 1 \tag{20}$$

ist

$$-\{3B_{0040}/(m+1)^2 + B_{0022} + 3(m+1)^2B_{0004}\} = \tfrac{1}{2}(14m^2+4m-1) > 0,$$
$$m = 1, \pm2, \pm3, \cdots$$

sowie

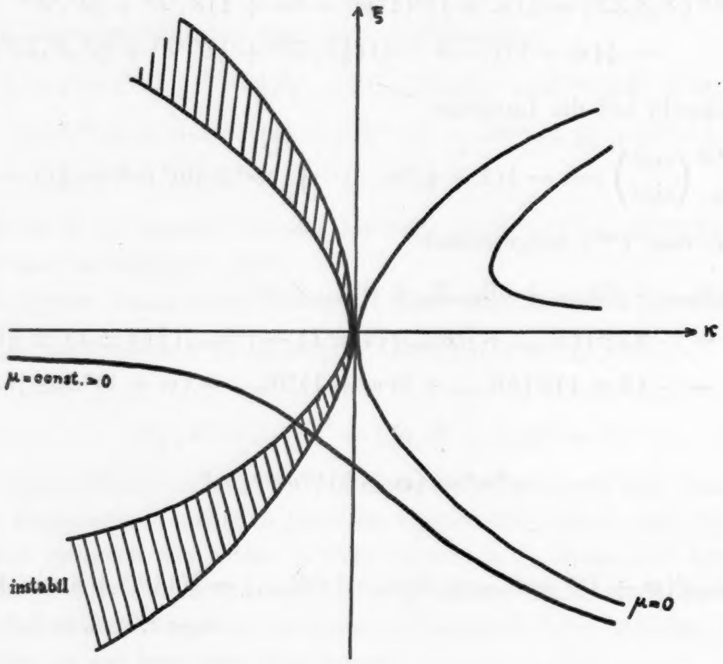
$$-2(B_{1020} + j^2B_{1002} + 3B_{0030}Z_0 + j^2B_{0012}Z_0)$$
$$= j(-3(3m+1) - (m+1) + (6m+2m+2)2m^2/a^2)$$
$$= 2(m+1)[(8m+2)m^2/a^2 - (5m+2)]$$
$$= 2(m+1)(2(4m+1)(m^2a^{-2}-1)+3m) > 0,$$
$$m = 1, \pm2, \pm3, \cdots.$$

Wir bekommen daher endgültig die beiden *parabelartigen Kurven*

$$\begin{aligned}\kappa &= -3\gamma_0\xi^2 + \cdots,\\ \kappa &= -\gamma_0\xi^2 + \cdots,\end{aligned} \qquad \gamma_0 = \frac{m+1}{4}\,\frac{14m^2+4m-1}{2(4m+1)(m^2a^{-2}-1)+3m} > 0;$$
$$m = 1, \pm2, \pm3, \cdots.$$

Sie sind beide in Richtung abnehmenden κ geöffnet. In dem (schraffierten) Gebiet *zwischen* ihnen sind die periodischen Bahnen einer Gruppe $\mu = \mu_0 = \text{const.}$ *instabil.*

Auch die halbganzen Hill'schen Periodenquotienten sind beim restringierten Dreikörperproblem zu beachten. Die Instabilitätsintervalle um diese Stellen herum sind von Wintner [2] untersucht worden. Hier liegt keine Verzweigung der periodischen Bahnen vor. Es gilt mit dem entsprechenden Zuwachs $\kappa = k - k^0$ der Jacobischen Konstante eine Entwicklung der Gestalt



$$\zeta = \kappa Z_0 + \mu Z_* + \cdots.$$

Das Instabilitätsgebiet besteht aus zwei vom Punkt $\kappa = 0$, $\mu = 0$ ausgehenden Winkelräumen $\mu > 0$ und $\mu < 0$, die in erster Annäherung durch zwei " Gerade "

$$\kappa = \gamma_1\mu + \cdots \text{ und } \kappa = \gamma_2\mu + \cdots$$

begrenzt werden.

UNIVERSITY OF LEIPZIG, GERMANY.

[2] A. Wintner, " On the periodic analytic continuations of the circular orbits in the restricted problem of three bodies," *National Academy of Sciences Proceedings*, vol. 22 (1936), pp. 435-439.

GEODESIC VERTICES ON SURFACES OF CONSTANT CURVATURE.*

By S. B. JACKSON.

1. Introduction. In a previous paper by the writer [6] [1] the attempt was made to characterize structurally, as far as possible, the closed plane curves of class C'' which have exactly two extrema of the curvature; i. e. two vertices. In that paper there were obtained five structural properties. The first part of the present paper is concerned with a discussion of the extent to which these properties can be carried over to characterize closed curves of class C'' which lie in a simply connected region of a surface of constant Gaussian curvature and which have only two geodesic vertices, i. e. extrema of the geodesic curvature. Four of these properties go over without alteration and one important new one is added (Theorem 6. 1) but the fifth one requires certain modifications (Theorems 8. 1 and 9. 1). The essential difference between the case of the surface and that of the plane lies in the geodesic circles which may have different structural properties from plane circles. In particular, they may have more than two points of intersection, a fact which is the basis for some of the examples given in the paper (**9** and **11**).

The last section of the paper (**11**) contains proofs for curves on surfaces of constant curvature of two theorems relating the number of geodesic vertices on a simple closed curve with its number of intersections with a geodesic circle (Theorems 11. 1 and 11. 2). These are direct extensions of known theorems on plane curves [6, Theorems 6. 1 and 7. 1] and are analogues of well known theorems on ovals [3, 1].

Scherk [9] has observed that by stereographic projection any theorem regarding vertices of plane curves is equivalent to one on the sphere concerning geodesic vertices. For the sphere such theorems as Theorem 5. 1, Theorem 11. 1, and Theorem 11. 2 follow trivially from the known results in the plane. Moreover, by using results of Mohrmann [8] regarding inflection points of curves on an ovaloid, these theorems may be readily obtained on the sphere directly, thence yielding simple proofs of the theorems in the

* Received March 16, 1948; revised July 19, 1949. Presented to the American Mathematical Society, September, 1947.

[1] The numbers in brackets refer to the bibliography.

plane.[2] It is interesting that these methods of Mohrmann were available in 1917, some years before the first explicit statements in the literature that every simple closed plane curve of class C'' has at least four vertices. As far as this writer is aware, the first statements of this theorem were by Fog [4] in 1933 and Graustein [5] in 1937, and neither of them used Mohrmann's methods. In the present paper however we consider curves in any simply connected region of a surface of constant curvature, without reference to its embedding in space. Mohrmann's methods do not appear directly applicable to this work therefore. Indeed, from one point of view this paper is an answer in one special case to a question raised by Scherk [9] as to what parts of these theorems can be salvaged in case the curves do not lie on an ovaloid.

2. Definitions and previous results. A *geodesic vertex* of an arc or curve of class C'' on a surface Σ is a point (or an arc of constant geodesic curvature) for which the geodesic curvature has a relative extremum with respect to the neighboring arcs. That is, if $1/\rho$ is the geodesic curvature at any point and $1/a$ the geodesic curvature at the vertex, then in a neighborhood of the vertex $1/\rho - 1/a$ does not change sign and is not identically zero. If, for two geodesic vertices, the geodesic curvatures are both relative maxima or both relative minima, they are said to be of the same type. For a plane curve, geodesic vertices coincide with ordinary vertices.

An arc on which the geodesic curvature is monotone non-increasing or monotone non-decreasing is called a *monotone arc.* If two monotone arcs both have the geodesic curvature non-increasing (non-decreasing) they are said to be of the same type, otherwise of opposite type. An arc or curve for which the geodesic curvature remains constant is called a *geodesic circle.*[3] A geodesic is a special case of a geodesic circle.

In discussing arcs lying on a surface it is often convenient to speak of one arc as lying locally to the right or to the left of another. Such a statement always implies the surface is viewed from the tip of the positive unit normal, and that the surface trihedral is right handed. An arc with positive geodesic curvature at a point thus lies locally to the left of the directed tangent geodesic at this point.

A simple closed arc of a curve which is never crossed by the remainder

[2] Cf. Scherk's review of [6] in *Mathematical Reviews,* vol. 6, p. 100, where he indicates how these theorems may be obtained.

[3] This is what Blaschke [2, § 7] calls a *Krümmungskreis,* as distinguished from an *Entfernungskreis,* which is the locus of points at a fixed geodesic distance from a given point.

of the curve is called a *simple loop*. If a simple closed curve, composed of differentiable arcs but allowing corners and lying in a simply connected region of a surface, is directed so that the simply connected region bounded by it lies to its left, we shall say that it is positively directed.

A transformation of class C'' of a region $\mathcal{S}$ on a surface Σ into the Gaussian plane is said to be of type I if (a) it is locally one to one, (b) it carries the geodesic circles of $\mathcal{S}$ into circles (or arcs of circles),[4] and (c) it preserves sense. The following results were obtained by the writer in an earlier paper [7, Theorems 3.1 and 3.2 and Lemma 4.7].

THEOREM 2.1. *A transformation of type I carries monotone arcs into monotone arcs of the same type. It carries geodesic vertices into geodesic vertices of the same type or into limit points of such vertices.*

THEOREM 2.2. *There is a transformation of type I, not necessarily one-to-one, taking any simply connected region $\mathcal{S}$ of a surface Σ of constant curvature into a region of the Gaussian plane.*

THEOREM 2.3. *If C is a positively directed simple closed curve, composed of a finite number of arcs of class C', which lies in a simply connected region $\mathcal{S}$ of a surface of constant curvature Σ, the image of C under any transformation of type I can contain no simple loop having no points of the curve lying to its left.*

As stated in the previous paper, Theorem 2.3 was for simple closed curves of class C'', but the proof holds without modification for the more general case.

In this paper we shall be concerned with arcs and curves having continuous geodesic curvature in a simply connected region $\mathcal{S}$ of a surface of constant curvature Σ, and this will be understood in all that follows except where it is indicated otherwise. The region $\mathcal{S}$ is assumed to contain no singularities of any kind. As it is used here, the term *arc* means the locally topological image of a line segment. This differs from the common topological use of the term in that the mapping need not be one-to-one in the large, so an arc may have double points. Similarly a *curve* is the locally topological image of a circle.

3. Geodesic circles. On a general surface, the locus of points at a given geodesic distance from a fixed point, or a distance circle as it is sometimes called, need not be a geodesic circle as defined above. However, it is

[4] In the Gaussian plane a line is a special case of a circle.

well known that for surfaces of constant curvature the distance circles are all geodesic circles [2, § 72]. It is not true however that all geodesic circles are distance circles. A distance circle which is a simple closed curve will be called a *complete geodesic circle.*

Although it is true, by Theorems 2. 1 and 2. 2, that geodesic circles in $\mathcal{S}$ go into circles[5] in the plane under transformations of type I, the fact that these transformations are not one-to-one makes it necessary to exercise caution about attributing to geodesic circles in $\mathcal{S}$ many of the familiar properties of circles in the plane. For example, it is not true that two geodesic circles can intersect in only two points. They can have any number of intersections, though under a transformation of type I all the intersections map into one or the other of the intersections of the corresponding plane circles. Examples of geodesic circles which intersect in more than two points will be given in **9** and **11**.

Let $\mathcal{R}$ denote the closed simply connected subregion of $\mathcal{S}$ bounded by any Jordan curve C in $\mathcal{S}$. Consider the family of complete geodesic circles contained in $\mathcal{R}$ with centers at a fixed arbitrary point P in $\mathcal{R}$, and let r be the least upper bound of the radii of these geodesic circles. Since $\mathcal{R}$ contains no boundary point of $\mathcal{S}$, it is clear that these geodesic circles are complete. If Σ has positive curvature, r must be less than the distance from P to a conjugate point P', for otherwise $\mathcal{S}$ would contain the region covered by all the geodesics through P and hence through P'. Since this region is isometric with a complete sphere, $\mathcal{S}$ would not be simply connected. If the curvature of Σ is non-positive there are no conjugate points in $\mathcal{R}$. The geodesic circle, O, with radius r and center P has at least one point in common with C, for if it did not, O and C, being closed sets would have a positive distance. A geodesic circle about P of larger radius would then belong to $\mathcal{R}$, which contradicts the definition of r. Clearly also no point of C lies inside O. From this and the discussion of conjugate points, it follows that every point in and on O is joined to P by a unique geodesic which is the path of minimum distance between these points.

If Q is any point on O and R is any point distinct from P on the geodesic segment PQ, the shortest path from R to any point of O is this geodesic segment RQ. For consider any other arc RS joining R to a point of O. We see that $PR + RQ = r = PS < PR + RS$ whence $RQ < RS$ which proves the contention. If, in particular, Q is a point where O meets C, RQ is the shortest path from any point R of PQ to the curve C. The results of this discussion may be summarized in the following statement.

[5] See footnote 4.

LEMMA 3.1. *If a Jordan curve C in $\mathcal{S}$ bounds the closed region $\mathcal{R}$, about any point P of $\mathcal{R}$ as center can be drawn a unique complete geodesic circle O contained in $\mathcal{R}$ and meeting C in one or more points. If Q is one of these common points and R is any point of the geodesic radius PQ, the minimum distance from R to C is along the geodesic segment RQ.*

LEMMA 3.2. *If $\mathcal{R}$ is the closed region bounded by a Jordan curve C in $\mathcal{S}$, and if C is divided in any manner into three arcs $\mathfrak{a}_1$, $\mathfrak{a}_2$, $\mathfrak{a}_3$, then there exists a complete geodesic circle contained in $\mathcal{R}$ and having points in common with all three arcs.*

In view of Lemma 3.1, the proof of Lemma 3.2 is an immediate adaptation of one suggested by Paul Erdös for the plane case [6, Lemma 3.1].

LEMMA 3.3. *If a Jordan curve C in $\mathcal{S}$ bounds the closed region $\mathcal{R}$, and P_0 is any point interior to $\mathcal{R}$, the largest complete geodesic circle in $\mathcal{R}$ which contains P_0 either in or on it has at least two points in common with C.*

It is clear that the radius of the largest geodesic circle in $\mathcal{R}$ with center P is a continuous function of P. Thus the set of points P for which P_0 belongs to the corresponding circle is compact, whence the maximum circle, O, mentioned in the lemma actually exists.

Lemma 3.3 may be made intuitively evident as follows. If O meets C only at a single point Q, a slight displacement of O yields a circle still containing P_0 but having no points in common with C. This contradicts the maximal property of O. A formal proof of the lemma can readily be constructed on these lines.

LEMMA 3.4. *A complete geodesic circle in $\mathcal{S}$ and its interior map in a one-to-one manner into the corresponding plane region under any transformation of type I.*

Under a transformation of type I a complete geodesic circle, being closed, maps into a plane circle traced one or more times. If the plane circle were traced more than once it would contain a simple loop without points of the curve to its left, which contradicts Theorem 2.3. The transformation is thus one-to-one on any complete geodesic circle, or, in fact on any two tangent complete geodesic circles since their images could meet only once. The lemma is now readily proved by showing any two points of the indicated region lie on such a pair of tangent complete geodesic circles.

Let O_1 represent any geodesic circle in $\mathcal{S}$, not necessarily complete. By a transformation of type I it goes into an arc of a circle, possibly over-

lapping itself. If O_1 has a double point, it must be a point of tangency, since this is true of the plane curve. O_1 is then merely a simple closed curve which maps one-to-one into its plane image, as in Lemma 3. 4. Applying Lemma 3. 2, we find that there is a complete geodesic circle O'_1 interior to O_1 and tangent to it at at least two distinct points. O_1 and O'_1 coincide near these tangencies since their plane images do, and are thus identical. We conclude that O_1 is a complete geodesic circle.

From this discussion, the facts about plane circles, and Lemma 3. 4 the following conclusion may be drawn.

LEMMA 3. 5. *A geodesic circle in $\mathcal{S}$ with a double point is a complete geodesic circle. Two geodesic circles, one of which is complete, either do not meet, or are tangent at just one point, or intersect at just two points.*

The lemma is false if the restriction that one circle be complete is removed. (Cf. **9**).

LEMMA 3. 6. *Every geodesic circle in $\mathcal{S}$ divides $\mathcal{S}$ into exactly two parts.*

If geodesic circle O has a double point, the conclusion follows from Lemma 3. 5 and the Jordan Curve Theorem, so it suffices to consider the case when O is a simple open arc.

A neighborhood of a point of $\mathcal{S}$ bounded by a complete geodesic circle about the point is called a complete circular neighborhood. The closure of such a neighborhood, by Lemma 3. 5, has at most a single arc in common with O, whence no point of $\mathcal{S} - O$ is a limit point of O. Moreover points of such a circular neighborhood which meets O may be classified as locally to the right or left of O. It follows readily that O divides $\mathcal{S}$ into at most two connected sets, namely the sets of points which can be joined in $\mathcal{S} - O$ to points lying locally to the right or left of O respectively.

It remains only to show that no two points P and Q locally on opposite sides of O can be joined by an arc in $\mathcal{S} - O$. Assume such an arc exists. It is clear that this arc may be completed into a simple closed curve by an arc PQ meeting O at a single point T. The region $\mathfrak{D}$ bounded by this curve can be covered, by the Heine-Borel Theorem, by a finite number of complete circular neighborhoods, the closure of each meeting O at most in a single arc. The subarc of O in $\mathfrak{D}$ is thus the sum of a finite number of closed arcs, whence it has an endpoint in $\mathfrak{D}$ since it meets the boundary only at T. Since O has no endpoint in $\mathcal{S}$ this is impossible and the lemma is proved.

As a result of Lemma 3. 6 it is possible to speak, in the large, of the subregion of $\mathcal{S}$ to the left or to the right of any directed geodesic circle.

LEMMA 3. 7. *A geodesic circle contained in a closed bounded region of $\mathcal{S}$ is complete.*

The region can be covered by a finite number of complete circular neighborhoods by the Heine-Borel Theorem. Each of these closed neighborhoods has at most a single arc in common with the given circle O, whence O consists of a finite number of closed arcs. Since, as above, O has no endpoint, this is possible only if O has a double point and is therefore complete by Lemma 3. 5.

4. Monotone arcs and geodesic vertices. The following two results characterizing geodesic vertices and monotone arcs are known to be true on any surface of sufficient differentiability [7, Corollary 2. 1, Lemmas 2. 1 and 2. 1].

LEMMA 4. 1. *A necessary and sufficient condition that an arc be monotone is that it cross every osculating geodesic circle at the point or arc of contact. The geodesic curvature is monotone non-decreasing or monotone non-increasing according as the crossing is from right to left or from left to right.*

LEMMA 4. 2. *In some neighborhood of a geodesic vertex, an arc lies entirely to one side of the osculating geodesic circle at this vertex, and conversely, a point (or arc) with this property is a geodesic vertex or a limit point of geodesic vertices. The geodesic curvature at the vertex is a maximum or a minimum according as the arc lies to the right or left of the osculating geodesic circle.*

In both cases the arc and the geodesic circle have locally only a single point or a single arc in common.

LEMMA 4. 3. *If P_0 and P_1 are any two points in that order on a monotone arc $\mathcal{A}$ in $\mathcal{S}$, the osculating geodesic circle at P_1 lies to the left or right of that at P_0 according as the geodesic curvature on $\mathcal{A}$ is non-decreasing or non-increasing. The two geodesic circles have no point in common unless they coincide and $\mathcal{A}$ contains the arc P_0P_1 of this geodesic circle.*

Let $\mathcal{S}$ be mapped into the plane by a transformation of type I, which is possible by Theorem 2. 2. Since geodesic circles go into circles and monotone arcs into monotone arcs under a transformation of type I, by Theorem 2. 1, the fact that the osculating geodesic circles at P_0 and P_1 meet only if they are identical and coincide with $\mathcal{A}$ from P_0 to P_1 follows from the corre-

sponding known result in the plane [6, Lemma 2. 5].[6] This implies immediately that once a monotone arc leaves its circle of curvature it never meets it again. The location of the geodesic circle at P_1 with reference to that at P_0 is then a consequence of Lemma 4. 1.

LEMMA 4. 4. *A monotone arc in $\mathcal{S}$ is simple except for any complete geodesic circles it contains. In this case the arc is tangent to itself without crossing at a single point or along a single arc of such a circle.*

This follows immediately by Lemma 4. 3. Since an arc without a geodesic vertex is monotone, Lemma 4. 4 leads at once to the following.

LEMMA 4. 5. *A simple closed arc in $\mathcal{S}$, not a geodesic circle, has at least one geodesic vertex interior to the arc.*

Let AB be a monotone arc of $\mathcal{S}$, not an arc of a geodesic circle, and let it be tangent to a geodesic circle O_1 at the point of minimum curvature, say B, and lie locally to the left of O_1 at B. Since AB lies locally to the left of O_1 at B, the same is true of its osculating geodesic circle at B. By Lemma 4. 3 A is definitely to the left of this geodesic circle and thus is not on O_1. A similar argument shows that if AB is tangent to and locally to the right of a geodesic circle O_2 at its point of maximum curvature, the other end is not on O_2. Taken together these statements establish the following result.

LEMMA 4. 6. *If an arc AB in $\mathcal{S}$, not an arc of a geodesic circle, is tangent to a geodesic circle in the same direction at A and B and lies locally on the same side of this circle at A and B, then there is a geodesic vertex interior to AB.*

We shall conclude this discussion with a result on plane arcs that will be useful in the next section.

LEMMA 4. 7. *Let a non-circular plane arc AB satisfy the following conditions:*

(a) *it is tangent to a circle (line) in the same direction at A and B;*

(b) *it lies locally to the left of this directed common tangent circle (line) at A and B;*

(c) *it contains no minimum of the curvature in its interior. Then AB*

---

[6] As was mentioned by Scherk in his review of [6] (*Mathematical Reviews*, vol. 6, p. 100) the work on plane monotone arcs is not new in the literature. See, for example, the papers by Vogt and Kneser there mentioned. As a matter of convenience however, references are given to [6] rather than to these original papers.

contains a simple loop having no points of the arc lying to its left. AB meets the circle (line) only at A and B, which may coincide.

Since, by a direct circular transformation, the given circle may be carried into a line, and all the properties are invariant under such transformations, it is enough to consider this case. By Lemma 4. 6, which certainly applies to the plane, AB is not monotone, and therefore has a maximum of curvature, which may be a point or an arc. Let M denote this point, or a point of this arc. As in the proof of Lemma 4. 6, the monotone arcs AM and MB lie entirely to the left of this common tangent line, so AB meets this line only at A and B. The arc AB is not simple since a simple arc satisfying the conditions contains a minimum [6, Cor. 4. 1. 1]. For the same reason it would not be simple even if we deleted from AB all the complete circles noted in Lemma 4. 3, so AM and MB meet. Since AM and MB are monotone and the curvature is non-negative at A and B by condition (b), it is positive interior to AB, and these arcs are respectively inwinding and outwinding spirals [6, Cor. 2. 5. 1 and 2. 5. 2]. The proof that the maximum of curvature lies on a simple loop with no points of AB to its left is identical with the proof that the maximum on a curve with two vertices lies on such a loop [6, Lemma 5. 4]. It will not be repeated here.

5. Location of geodesic vertices.

LEMMA 5. 1. *If a simple arc AB in $\mathcal{S}$, not an arc of a geodesic circle, is tangent to a complete geodesic circle O in the same direction at A and B and never crosses this circle, there is at least one minimum or at least one maximum of the geodesic curvature interior to AB according as AB lies to the left or to the right of this geodesic circle. A and B may coincide.*

If the arc AB lies interior to O, the lemma follows from the one-to-one mapping of type I guaranteed by Lemma 3. 4 and the corresponding known result in the plane [6, Lemma 4. 1]. Let AB then lie outside O. The arc AB may be completed into a simple closed curve C of class C' by adjoining the directed circular arc BA. Let $\mathfrak{R}$ denote the closed region bounded by C. It is sufficient to prove the lemma when AB is so directed that C is a positively directed curve.

Consider first the case when O is exterior to $\mathfrak{R}$, whence AB lies to the left of O. Let C be mapped by a transformation of type I into the plane curve C. Arc AB contains a minimum of the geodesic curvature since otherwise Lemma 4. 7 and Theorem 2. 3 yield contradictory results on its plane image, which establishes the lemma in this case.

In the contrary case O is contained in $\mathfrak{R}$ and the arc AB lies to the right of O. It remains only to show the existence of a maximum on AB. By Lemma 4. 6, AB is not monotone, and it is clearly sufficient to consider the case when there is just one geodesic vertex, a point of which we denote by E. Let RS be a subarc of AB containing E and contained in a complete geodesic circle O about E. Consider the complete geodesic circle O'' guaranteed in Lemma 3. 2 meeting arcs RE, ES, and $SBAR$ of C. The points of contact are necessarily tangencies. Since O'' clearly cannot meet circular arc BA, by Lemma 4. 6 it contains E only if it contains arc RS, in which case this arc is a maximum by Lemma 4. 2. If E is not on O'' a subarc of RS is tangent to O'' at two points and the existence of a maximum is assured by mapping O' by a transformation of type I and using Theorem 2. 1 and the known results in the plane [6, Lemma 4. 1]. This completes the proof of the lemma.

The restriction that AB shall not cross O in Lemma 5. 1 may be lightened as follows.

LEMMA 5. 2. *If a simple arc AB in $\mathcal{S}$, not an arc of a geodesic circle, is tangent to a complete geodesic circle O in the same direction at A and B and lies locally on the same side of O at A and B, there is at least one minimum or at least one maximum of the geodesic curvature interior to AB according as AB lies locally to the left or right of O at these points. A and B may coincide.*

By Lemma 4. 6 there is at least one geodesic vertex interior to AB. Consider the case when AB lies locally to the right of O and assume the lemma is false, i. e. assume the only geodesic vertex is a minimum. Since O lies to the left of AB at A and B, it lies to the left of the osculating geodesic circles at A and B. If the only vertex is a minimum, then by Lemma 4. 3 O lies to the left of all the osculating geodesic circles. Arc AB thus never crosses O and lies entirely to its right. But Lemma 5. 1 then states that AB has a maximum of geodesic curvature. The contradiction proves the lemma for this case. The remaining case when AB is locally to the left of O may be reduced to the one above by reversing the sense on AB. This completes the proof.

Lemma 5. 1 leads easily to the following useful result.

LEMMA 5. 3. *Let AB be a simple closed arc in $\mathcal{S}$, not a geodesic circle. Let $\mathfrak{R}$ denote the region bounded by AB, and θ the positive angle interior to $\mathfrak{R}$ between the two tangents at the double point. If $\theta \geqq \pi$ there is a maximum or a minimum of the geodesic curvature interior to AB according*

as $\mathfrak{R}$ lies to the right or left of AB. If $\theta \leqq \pi$ there is a maximum or a minimum of the geodesic curvature interior to AB according as $\mathfrak{R}$ lies to the left or right of AB.

It is clearly possible to draw an arbitrarily small complete geodesic circle tangent to AB at points A' and B' near A and B. It may be that these four points will coincide. This complete geodesic circle can be drawn exterior to $\mathfrak{R}$ or interior to $\mathfrak{R}$ according as $\theta \geqq \pi$ or $\theta \leqq \pi$. The proof is then immediate from Lemma 5. 1.

It is interesting to observe that Lemma 5. 3 remains valid when $\mathfrak{R}$ is thought of as the region *exterior* to AB in $\mathfrak{S}$. This follows easily from the lemma itself.

Lemma 5. 3 leads trivially to one of the main results of a previous paper [7, Theorem 4. 1], namely the Four-vertex Theorem for $\mathfrak{S}$.

THEOREM 5. 1. *Every simple closed curve of class C'', not a geodesic circle, in a simply connected region $\mathfrak{S}$ of a surface of constant curvature has at least four geodesic vertices.*

Let the curve be positively directed, and let M be a geodesic vertex. Consider the arc from M to M as the arc AB of Lemma 5. 3. Since $\theta = \pi$ the arc contains both a maximum and a minimum distinct from M. This proves the theorem since the number of geodesic vertices is even if it is finite.

6. Curves with two geodesic vertices. We shall turn our attention to obtaining structural properties of curves in $\mathfrak{S}$ having exactly two geodesic vertices. Such a curve, C, consists of two monotone arcs of opposite type. As was noted in Lemma 4. 4, these monotone arcs may contain complete geodesic circles, and one or both of the geodesic vertices may be such circles. The curve which is obtained from C by removing all such complete geodesic circles is called the normalized curve, $\bar{C}$, corresponding to C. $\bar{C}$ is still of class C'', and the process neither gains nor loses geodesic vertices.[7] In the following discussion $\bar{C}$ always denotes the normalized curve corresponding to C.

A double point is called simple if the curve passes through the point only twice.

We proceed to establish the following five properties of C.

THEOREM 6. 1. *If C is a curve of class C'', not a geodesic circle, in a simply connected region $\mathfrak{S}$ of a surface of constant curvature, and if C has exactly two geodesic vertices, then:*

---

[7] If a complete geodesic circle is traced completely k times and partly so again, it is understood that in $\bar{C}$ the k complete revolutions are omitted, but the remaining arc left, so $\bar{C}$ is closed and of class C''.

(a) *the normalized curve $\bar{C}$ may be divided into two simple arcs;*

(b) *the normalized curve $\bar{C}$ has double points, but all of them are simple;*

(c) *at any point of tangency of C with itself the directed tangents coincide;*

(d) *C contains exactly two simple loops, one loop containing the maximum of geodesic curvature and having no points of C to its left, the other containing the minimum of geodesic curvature and having no points of C to its right;*

(e) *C has only a finite number of double points and double arcs.*

Property (a) follows from Lemma 4.4 and the fact that the two monotone arcs making up $\bar{C}$ are simple by construction.

In (b), the fact the $\bar{C}$ has double points is precisely the content of Theorem 5.1. If P were a double point of $\bar{C}$ which was not simple, it would divide $\bar{C}$ into at least three closed arcs, none of which are geodesic circles by construction of $\bar{C}$. By Lemma 4.5 each of these arcs would have a geodesic vertex, contradicting the hypothesis of only two geodesic vertices on C. Property (b) can also be proved directly from Lemma 4.4, since neither of the two monotone arcs can pass through any point but once.

To prove (c) let $\mathcal{S}$ be mapped into the plane by a transformation of type I. The result then follows from the corresponding property for plane curves with just two vertices [6, Theorem 5.1].

For the proof of (d), let m and M be the points of minimum and maximum geodesic curvature respectively (or points on the arcs of minimum and maximum geodesic curvature). If the maximum of geodesic curvature is a complete geodesic circle, this circle is itself the required simple loop since, by Lemma 4.3, no point of C lies to its left. In the contrary case, by Lemma 4.4, arcs mM and Mm are simple except for any complete geodesic circles they may contain. Moreover they intersect, for otherwise C is simple, contrary to (b). Let A be the first point where Mm meets mM. We shall show that AMA is the required simple loop. The two arcs are not tangent at A, for the directed tangents would coincide by (c), whence the arc would contain both a maximum and a minimum by Lemma 5.3, and this is impossible. Since M is a maximum, by Lemma 5.3 the arcs pass at A into the region to the right of AMA. It should be noted that this region may be the exterior of the arc. By construction, arc mA does not cross AMA and thus lies to its right. If AMA is not the required simple loop the arc Am must cross it. Let B be the first such crossing. Since Mm never crosses itself, B must lie on AM, and if B is the first crossing Am approaches

AM from the right at B. Consider arc $BMAB$. If the region $\mathfrak{R}$ bounded by it (inside it) lies to its left the angle θ interior to $\mathfrak{R}$ between the tangents at B is greater than or equal to π. If $\mathfrak{R}$ lies to the right of $BMAB$, then $\theta \leqq \pi$. In either case there would be a minimum on $BMAB$ by Lemma 5. 3. Since this is false, the point B does not exist and AMA is the required simple loop. The case of the minimum of geodesic curvature is reduced to this by reversing the direction on C, which completes the proof of (d).

Finally, consider property (e). Whenever one of the two monotone arcs meets itself, it contains a complete geodesic circle by Lemma 4. 4. These double points and arcs are thus surely finite in number, and hence isolated on the arc. At a point where the two monotone arcs are tangent to each other but have different geodesic curvatures, the directed tangents coincide, by (c), and it is readily verified that the arcs have locally only the single point of tangency in common, so that such tangencies are isolated double points. At a point or arc of tangency where the geodesic curvatures are equal, the osculating geodesic circles coincide. Indeed the arc of contact, if any, is an arc of this circle since one arc has monotone non-decreasing and the other monotone non-decreasing geodesic curvature. But by Lemma 4. 1 one of the arcs crosses this circle from right to left and the other from left to right. The osculating geodesic circle thus acts as a barrier and the two arcs have no other points in common in some neighborhood of this point or arc of contact which is thus isolated. If C had an infinite number of double points and arcs, they would have a limit point, which would necessarily be a tangency. But since all the tangencies have been shown to be isolated this is impossible, and property (e) is established. Properties (a), (b), (c), (d) above are already known for plane curves with two vertices [6, Theorem 5. 1], while property (e) is new.

By (e) the curve C divides $\mathcal{S}$ into a finite number of regions, which are all simply connected except one, called the exterior.[8] Two of the regions, by (d), are completely bounded by the simple loops. For plane curves it is known [6, Theorem 5. 1] that *no region determined by C except the two indicated in* (d) *is bounded in the same sense by all the arcs of C which bound it.*

The next three sections of the paper are devoted to a discussion of the validity of this statement for curves in $\mathcal{S}$.

---

[8] In the case of the Gaussian plane, or equivalently the sphere, there is no real distinction between interior and exterior regions, since any finite point can be carried into the point at infinity by a direct circular transformation. In general, however, the distinction between exterior and interior regions is genuine, as will be made clear in the work that follows.

7. Angular measure of transformed curves. Let C be a positively directed simple closed curve of class[9] D' in $\mathcal{S}$, and let K be the plane image of C under a transformation of type I which carries the interior of C into a finite region of the plane. Such a transformation always exists unless every point of the plane is the image of some point in this interior. K is also of class D'. Let C be deformed, as is always possible, through simple closed curves of class D' lying in its interior into a simple closed curve C_0 lying in a complete geodesic circle of $\mathcal{S}$, and let the deformation be such that at any corner the angle between the directed tangents remains in the open interval $(-\pi, \pi)$. Curve K is thereby similarly deformed through finite points into a curve K_0, which is a positively directed simple closed curve by Lemma 3.4.

The angular measure of a closed plane curve of class D' is defined as the total rotation of a directed tangent on tracing the curve once. At a corner this includes the directed angle in the interval $(-\pi, \pi)$ through which the first directed tangent must turn to coincide with the second. This is equivalent to rounding off each corner with a small circular arc and considering the angular measure of the resulting curve of class C'. Since K_0 is a simple closed curve, its angular measure is 2π, [10]. We conclude K also has angular measure 2π since the angular measure varies continuously (the deformation being through finite points) but is an integral multiple of 2π. This justifies the following statement.

LEMMA 7.1. *The image, under a transformation of type I, of any positively directed simple closed curve of class D' in $\mathcal{S}$ whose interior has a finite image has an angular measure of 2π.*

One further result regarding the angular measure of plane curves will be convenient.

LEMMA 7.2. *Let K be a closed plane curve of class D' for which there is a point O not on K, satisfying the following condition: the angle θ from some fixed direction to OP changes monotonically as P traces K. Then the angular measure of K equals the total variation in θ.*

It is sufficient to prove the result when θ is monotone increasing. At a corner of K, the various positions assumed by a directed line rotating from the first directed tangent to the second will also be called tangents, for convenience. Consider a point P of K and a directed tangent at P. The angle θ from the given fixed direction to OP, the angle λ from this same fixed

[9] A curve or arc is said to be of class D' if it is a finite succession of differentiable arcs. It is of class C' except for a finite number of corners.

direction to the directed tangent, and the angle ϕ from OP to the directed tangent are clearly related by the equation $\lambda = \phi + \theta$, whence it follows that $\triangle\lambda = \triangle\phi + \triangle\theta$ where the $\triangle$ indicates the total variation around K. Since θ is increasing it follows that $0 < \phi < \pi$. But since $\triangle\phi$ is an integral multiple of 2π, it follows that $\triangle\phi = 0$. This proves the lemma, for $\triangle\lambda$ is the angular measure of K.

8. Arcs bounding the exterior region.

LEMMA 8. 1. *Let AB and BA be two simple monotone arcs of opposite type in $\mathcal{S}$, meeting each other only at A and B and not tangent to each other in opposite directions at these points. Then there exists a transformation of type I of the closed region $\mathcal{R}$ bounded by the two arcs which is one-to-one and which carries this region into a finite region of the plane.*

Let the arcs be directed so that the region bounded by them lies to their left. Since the two arcs are of opposite type, one of the ends, say A, is the point of minimum geodesic curvature on both arcs.

Suppose first that the angle interior to $\mathcal{R}$ at A is less than or equal to π, and let P be any point interior to $\mathcal{R}$. By Lemma 3. 3 there is a complete geodesic circle containing P which is contained in $\mathcal{R}$ and has at least two points in common with its boundary. Since B is the only point where this circle could meet the boundary and not be tangent to it, this circle is tangent to one of the two arcs. There exists a complete geodesic circle passing through P which is contained in the one just constructed and has the same point of contact with one of the given arcs. This fact follows from the possibility of the construction in the plane and the fact that a transformation of type I on the complete geodesic circle is one-to-one. The point P therefore lies on a complete geodesic circle tangent to one of the arcs and lying to the left of both the arc and its osculating geodesic circle at the point of contact.

If, on the other hand, the angle at A in $\mathcal{R}$ exceeds π, draw the osculating geodesic circle to BA at A, which clearly passes to the interior of $\mathcal{R}$ at A. Since by Lemma 3. 6, it divides $\mathcal{R}$ into two parts, it meets the boundary again for the first time in a point Q. The point Q must be on arc AB, for BA has no further points in common with this circular arc unless BA is a geodesic circle, in which case Q coincides with B. This circular arc, AQ, divides $\mathcal{R}$ into the subregions $\mathcal{R}_1$ and $\mathcal{R}_2$ where $\mathcal{R}_1$ denotes the subregion containing arc BA in its boundary. Since the only angle interior to $\mathcal{R}$ which may exceed π is at B we apply Lemma 3. 3 as before to any point P of $\mathcal{R}_1$, and again conclude that P lies on a complete geodesic circle which

is tangent to and to the left of one of the osculating geodesic circles of the given arcs. If P belongs to $\mathfrak{R}_2$, Lemma 3. 3 assures us of a complete geodesic circle containing P and meeting the boundary of $\mathfrak{R}_2$ at least twice. The angles interior to $\mathfrak{R}_2$ at A and Q are acute, so both contacts are tangencies. Moreover, by Lemma 3. 5 a geodesic circle cannot be tangent to a complete geodesic circle at two points, so one of these tangencies is on the given arc AQ. The argument used above shows as before that P lies on a complete geodesic circle tangent to and to the left of an osculating geodesic circle of one of the given arcs. This property has now been established for all points in $\mathfrak{R}$.

Since the arcs do not have oppositely directed common tangents at A, there exists a point of $\mathcal{S}$ lying to the right of the two osculating geodesic circles at A and contained in a complete geodesic circle about A. Consider any transformation of type I of $\mathcal{S}$ which carries this point into the point at infinity in the plane [cf. 7, Lemma 3. 5]. The two osculating geodesic circles at A are thereby carried into plane circles whose interiors lie to their left. That is so say, the transformed arcs $A'B'$ and $B'A'$ both have positive curvature at A', and since at A' both arcs have their minimum curvature, it follows that both arcs have positive curvature at every point. All osculating circles lie interior to one or the other of the osculating circles at A' by Lemma 4. 3, and all these osculating circles have their interiors to their left. Since we have shown that every point P in $\mathfrak{R}$ lies on a complete geodesic circle tangent to and to the left of some osculating geodesic circle of AB or BA, the image point P' lies on a circle tangent to and inside of one of the osculating circles of $A'B'$ or $B'A'$. This shows that the transformation takes every point P of $\mathfrak{R}$ into a finite point. In fact the entire region $\mathfrak{R}$ is mapped into the sum of the two osculating circles at A'.

By Lemma 7. 1 the curve of class D' consisting of the arcs $A'B'$ and $B'A'$ has an angular measure of 2π. Let this curve be taken as the curve K of Lemma 7. 2. Any point lying to the left of both the osculating circles at B satisfies the conditions of the point O of the lemma, since it lies interior to all the osculating circles, so the vector OP always turns in the positive sense as P traces K. By Lemma 7. 2 $\triangle\theta = 2\pi$; i. e. the radius vector OP turns around exactly once, turning always in the positive sense. This means that K is a positively directed simple closed curve. The transformation from the curve in S to K is then one-to-one. It is known that the interior of K is the topological image of one or more simply connected subregions of $\mathfrak{R}$ [7, Lemma 4. 5]. However, since the boundary of $\mathfrak{R}$ maps into K one-to-one, the points of $\mathfrak{R}$ near its boundary map into the points in the plane close to K and interior to it. Since this set of points of R is

connected, they all belong to the same component of the inverse image of the interior of K. But the only simply connected subregion of $\mathcal{R}$ containing all these points is the entire interior of $\mathcal{R}$. The transformation is thus proved to be one-to-one on $\mathcal{R}$, which completes the proof of the lemma.

Lemma 8.1 gives rise easily to the facts stated in the next two lemmas.

LEMMA 8.2. *If C is a curve in $\mathcal{S}$ having just two geodesic vertices, and if the exterior region is bounded entirely by one of the simple loops of C, then none of the regions of $\mathcal{S}$ determined by C is bounded in the same sense by all of its bounding arcs except the two bounded entirely by the simple loops.*

LEMMA 8.3. *If C is a curve in $\mathcal{S}$ having just two geodesic vertices, it is not possible for the boundary of the exterior region to consist of exactly two arcs which bound this region in the same sense.*

For Lemma 8.2 the arcs AB and BA are the arcs into which the vertex divides the simple loop, while for Lemma 8.3 they are the two alleged bounding arcs. In both cases Lemma 8.1 guarantees a one-to-one transformation of type I of the interiors of these arcs, from which the conclusions follow from the known facts for plane curves with two vertices [6, Theorem 5.1 (d); see also Section **6** above].

We turn finally to the possibility that the boundary of the exterior region for a curve C in $\mathcal{S}$ having just two geodesic vertices shall consist of more than two arcs, all bounding it in the same sense. Assume this to be the case, and let C be so directed that the exterior lies to the right of its bounding arcs. It can be easily verified by use of Theorem 6.1 that this sequence of arcs, finite in number, form a simple closed curve which is the positively directed boundary of a simply connected region $\mathcal{R}$.

The arcs of monotone increasing and decreasing geodesic curvature constituting C will be denoted by A_i and A_d, and their subarcs called i-arcs and d-arcs respectively. Since neither A_i or A_d crosses itself, it is clear that the i-arcs and d-arcs alternate on the boundary of $\mathcal{R}$. Thus if P_1 denotes the point of the boundary of $\mathcal{R}$ nearest the minimum of geodesic curvature on A_i and if succeeding endpoints taken in order about the boundary of $\mathcal{R}$ are $Q_1, P_2, Q_2, \cdots, P_n, Q_n$, arcs P_kQ_k are i-arcs while arcs Q_kP_{k-1} are d-arcs. It is assumed that all subscripts are reduced mod n, and by Lemma 8.3 $n \geqq 2$.

It is relatively simple deduction, using Lemma 4.4, that the order in which these endpoints occur on A_i and A_d coincides with the cyclic order, given above, in which they occur on the boundary of $\mathcal{R}$. The first of these points on A_i is P_1, while the first on A_d is, say, Q_s. For every pair of consecutive endpoints except P_sQ_s and Q_nP_1 it follows that there is both an

i-arc and a d-arc joining them. If P is any point of $\mathfrak{R}$ between such a pair of arcs, Lemma 3. 3 guarantees the existence of a geodesic circle containing P and tangent to both arcs, since neither arc can be tangent to the circle twice by Lemma 4. 6. As in Lemma 8. 1 we can thus obtain a geodesic circle passing through P, tangent to the boundary of $\mathfrak{R}$ and lying to the left of the osculating geodesic circle at the point of contact.

Any other point of R lies to the left of all the d-arcs and i-arcs here mentioned. Since a discussion of the various possible cases, namely $s = 1$, $s = n$, and $s \neq 1, n$, reveals that the subregion of $\mathfrak{R}$ bounded by these arcs and containing P can contain at most one interior angle exceeding π, it follows as above using Lemma 3. 3 that there is a complete geodesic circle passing through P, tangent to A_i or A_d and lying to the left of its osculating geodesic circle at the point of contact. This property has now been established for all points of $\mathfrak{R}$.

Let $\mathcal{S}$ be mapped into the plane by a transformation of type I taking some point to the right of the osculating geodesic circle of C at the point of minimum geodesic curvature into the point at infinity in the plane. The curvature of the transformed curve K is then always positive, and all osculating circles lie interior to and to the left of the one of minimum curvature, by Lemma 4. 3. By use of Lemma 7. 1 and 7. 2, taking O as any point in the smallest osculating circle of K, it follows exactly as in Lemma 8. 1 that the transformation maps $\mathfrak{R}$ one-to-one into a finite portion of the plane. This would mean that the exterior region for the plane curve K has a boundary which consists of more than two arcs all bounding it in the same sense, which is known to be impossible [6, Theorem 5. 1 (d)]. It is therefore impossible for the exterior region of C on $\mathcal{S}$ to be bounded in the same sense by more than two arcs. This result and the facts stated in Lemmas 8. 2 and 8. 3 may be collected in the following theorem.

THEOREM 8. 1. *If C is a curve of class C'', not a geodesic circle, in a simply connected region $\mathcal{S}$ of a surface of constant curvature, and if C has exactly two geodesic vertices, then the exterior region of C on $\mathcal{S}$ can be bounded in the same sense by all its bounding arcs only if it is one of the two regions completely bounded by a simple loop. In this case, no region determined by C except these two is bounded in the same sense by all its bounding arcs.*

9. Arcs bounding an interior region. Theorem 8. 1 answers affirmatively for the exterior region the question as to whether the conjecture at the end of section **6** is valid on $\mathcal{S}$. It remains to consider the interior regions.

LEMMA 9.1. *If a curve C in $\mathcal{S}$ has just two geodesic vertices, no interior region determined by C whose boundary consists of more than two arcs can be bounded in the same sense by all these bounding arcs.*

Suppose such a region $\mathcal{R}$ exists. Let any two adjacent arcs of its boundary be called $\mathcal{A}_1$ and $\mathcal{A}_2$, and let the remainder of the boundary be called $\mathcal{A}_3$. By Lemma 3. 2 there is a complete geodesic circle in $\mathcal{R}$ having points in common with all three arcs. Since all angles interior to $\mathcal{R}$ at its corners do not exceed π, these contacts are all similarly directed tangencies. The tangencies divide C into at least three arcs, each of which either contains a geodesic vertex by Lemma 4. 6 or is a geodesic vertex by Lemma 4. 2. This contradicts the hypothesis that C has only two geodesic vertices, whence no such region $\mathcal{R}$ can exist.

The conjecture of section **6** would be completely proved if we could show finally that no interior region could be bounded in the same sense by just two arcs. Strangely enough, this is false, as will be demonstrated by an example.

It is readily shown that the lemniscate, whose polar equation is $r^2 = \operatorname{Cos} 2\theta$, has just two vertices, and that the inflectional tangents at the double point are the osculating circles there. The curve is directed so that the right hand loop is positively traced, inducing a direction also on the inflectional tangents. Certain points on the curve and its tangents will be denoted as follows: $P'(0,0)$, $A'(1,0)$, $B'(-1,-1)$, $D'(1,1)$, $E'(-1,0)$, $F'(1,-1)$, $G'(-1,1)$, and C' the point at infinity, the given coordinates being rectangular. Let us subject the figure to the complex linear transformation $w = (1-iz)/(z-i)$ and denote the transforms of $A', B', \cdots,$ by $A, B, \cdots$. In the transformed figure the origin O is exterior to both the loops PAP and PEP but is interior to the two osculating circles at P, which are oppositely directed. Consider the curve $PAPBCDPEPFCGP$ obtained by inserting the osculating circles into the original curve. Since the transformation is of type I the only vertices on this curve are at A and E. Instead of considering this as a curve in the plane let us consider it on the logarithmic Riemann surface with branch point at O, i. e. the Riemann surface for $w = \log z$.

Tracing out the loop PAP brings us back to the same point P of this surface, since O is not contained in this loop, but when we trace the first osculating circle, that does contain O, and therefore leads to a point $\bar{P}$ corresponding to P but on a different sheet of the surface. Upon tracing the second loop, not containing O, we return to $\bar{P}$, but on tracing the other osculating circle, O is traversed in the reverse direction and we therefore

return to P. The curve is therefore a closed curve on the Riemann surface. Moreover the osculating circles clearly have a common point C on the surface. On the surface the curve can thus be described as the curve $PAPBCD\bar{P}E\bar{P}FCGP$. The only double points are P, $\bar{P}$, and C, and they are all simple. It is immediately clear that there are two interior regions of this curve which are bounded in the same sense by the two arcs which bound them. They are the region bounded by circular arcs $CD\bar{P}$ and $\bar{P}FC$ and the region bounded by circular arcs PBC and CGP.

Considering the curve as a curve of the Riemann surface was done merely for convenience in the discussion. The configuration can occur on an ordinary developable surface. For example, the conical surface with vertex at the origin and a circular helix about the z-axis as directing curve will serve the purpose, as will the tangent surface of a circular helix. The conjecture of section **6** is thus false for the interior regions of the curve C. It should be emphasized that while, as here, the surface Σ may contain branch points or other singularities, region $\mathcal{S}$ is assumed free of them. Here $\mathcal{S}$ may consist of that part of the Riemann surface covering an annular ring about O. This region is simply connected and free of singularities. Similar remarks apply to the regions $\mathcal{S}$ in the examples of section **11**. The results for interior regions may be summarized in the following theorem.

THEOREM 9.1. *If C is a curve of class C'', not a geodesic circle, in a simply connected region $\mathcal{S}$ of a surface Σ of constant curvature, and if C has exactly two geodesic vertices, any interior region determined by C and bounded in the same sense by all its bounding arcs is either completely bounded by one of the simple loops, or is bounded by exactly two arcs. If regions of this latter type occur, the regions bounded by the simple loops are both interior regions.*

It is interesting to observe that in the example just given the critical point is that the two circles intersect in three distinct points, namely P, $\bar{P}$, and C. By continuing the subarcs so that their plane images are traced n times, these arcs meet in $2n - 1$ points and the number of regions bounded as above by two arcs is $2n$. Thus we can construct examples in which the number of regions bounded by two arcs in the same sense is arbitrarily large. As far as the circles themselves are concerned, it should be noted that in the simply connected region consisting of the part of the surface covering a circular ring in the plane about the origin, the circles intersect infinitely many times. Similarly we can construct examples of circles in such a simply connected region which are tangent at infinitely many distinct points.

10. Counterexample. As has already been noted in Lemma 4. 3, monotone arcs have an essentially spiral character. In fact any plane curve with just two vertices can be reduced by a direct circular transformation to a curve in which one of the two monotone arcs is an inwinding spiral, and the other an outwinding spiral [6, p. 573]. For convenience we will suppose the curves normalized, so the only double points are where the two monotone arcs meet each other. The following question then naturally arises. Let the double points or arcs be arranged in the order $P_1, P_2, \cdots, P_n$ in which they occur on one of the monotone arcs. Will they occur in the reverse order $P_n, P_{n-1}, \cdots, P_1$ on the other arc? The same question can be phrased in a different but equivalent way as follows. Is every double point or arc a cut point or arc for the set of points making up the curve C? Simple examples show that this is often the case. If the question could be answered in the affirmative, it would be a very strong structural characteristic of curves with just two vertices. Actually the answer is in the negative, however, as is shown by the following example of a normalized curve in the plane having five double points, for which the order of the points on one arc is P_1, P_2, P_3, P_4, P_5, while on the other arc the order is P_5, P_2, P_3, P_4, P_1.

A curve is determined by the following equations if the radius of curvature R is given as a function of the slope angle ϕ

$$x = x_0 + \int_0^\phi R \operatorname{Cos} \phi \, d\phi \qquad y = y_0 + \int_0^\phi R \operatorname{Sin} \phi \, d\phi.$$

Let $x_0 = y_0 = 0$, and consider the curve (Figure 1) defined as follows for $|\phi| \leqq 4\pi$.

$$R = \begin{cases} 1 + |\phi|, & 0 \leqq |\phi| \leqq \pi \\ 1 + \pi, & \pi \leqq |\phi| \leqq 2\pi \\ 1 + |\phi| - \pi, & 2\pi \leqq |\phi| \leqq 3\pi \\ 1 + 2|\phi| - 4\pi, & 3\pi \leqq |\phi| \leqq 4\pi. \end{cases}$$

As defined above, R is positive and continuous, and the curve consists of two monotone arcs. The arc $-4\pi \leqq \phi \leqq 0$ is an arc of non-decreasing curvature, while the arc $0 \leqq \phi \leqq 4\pi$ is an arc of non-increasing curvature. It is a routine matter to carry out the computation to find the various points where the tangents are horizontal.

These points, together with the corresponding values of ϕ are as follows.

$\phi = 0,$	$O(0,0)$		
$\phi = \pi$	$A(-2, 2+\pi)$;	$\phi = -\pi$	$A'(2, 2+\pi)$
$\phi = 2\pi$	$B(-2, -\pi)$;	$\phi = -2\pi$	$B'(2, -\pi)$
$\phi = 3\pi$	$C(-4, 2+2\pi)$;	$\phi = -3\pi$	$C'(4, 2+2\pi)$
$\phi = 4\pi$ and -4π,		$D(0, -4\pi)$.	

Since at D the tangents and curvatures of the two arcs coincide, it is clear that this is actually a closed curve of class C'', having just two vertices, O and D, and with angular measure 8π.

Since, by its definition, the curve is symmetric in the y-axis, all the points where the curve meets this axis are double points. P_1 is the first point where arc OD meets the y-axis, and thus, by symmetry, the first double point on this arc. The open subarc P_1AB clearly does not meet the y-axis, but it is readily found that BC crosses the axis twice since B and C have negative x-coordinates, while the point where $\phi = 5\pi/2$ has a positive value for x. These two double points are denoted by P_3 and P_5. However, it

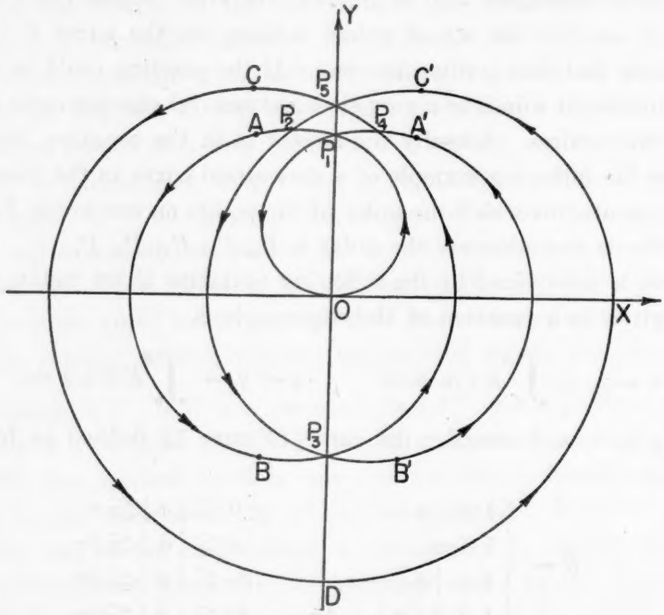


FIGURE 1.

appears at once that arc OD has positive slope at P_3, while DO has negative slope there. The crossing is therefore in the direction indicated, proving the existence of the other two symmetric double points P_2 and P_4 on arcs P_1P_3 and P_3P_1 respectively. The double points therefore occur on the monotone arcs in the orders indicated above. This answers in the negative the question asked at the beginning of this section.

11. Geodesic vertices on simple closed curves. Closely related to the classical four-vertex theorem are several other theorems relating the number of vertices on an oval to the number of its intersections with a circle [1, 2 p. 49, 3]. These theorems can be extended to simple closed plane curves

[6, § 6 and § 7]. It is the purpose of this section to investigate the extent to which these theorems can be generalized to simple closed curves on $\mathcal{S}$.

THEOREM 11. 1. *A simple closed curve C of class C'' in $\mathcal{S}$ not a geodesic circle which meets any geodesic circle at most four times has exactly four geodesic vertices.*

Let C be positively directed and consider a point or arc M of maximum geodesic curvature. We will show that the osculating geodesic circle O at M is a complete geodesic circle lying in the closed region $\mathfrak{R}$ bounded by C and meeting C only at M. The osculating geodesic circle O at M lies locally interior to C near M by Lemma 4. 2. If it is not contained in R, both branches of the circle from M must cross C, since by Lemma 3. 6 each arc of the circle divides R into two parts. It will then be possible to select points A, B, M, B', A' in that order on O with A, A' exterior to C and B, B' interior to C. Let O be deformed into an arbitrarily near geodesic circle by decreasing the geodesic curvature slightly, preserving the tangency with C at M (or some point of M). The points A, A', B, B' will vary continuously into points A_1, A'_1, B_1, B'_1 on the new circle O_1, but the deformation may be taken so small that these points do not cross C. The new circle lies locally to the right of C at M by Lemma 4. 2, so it is possible to choose points P, P' of O_1 exterior to C on arcs MB_1 and MB'_1 respectively. Thus, in addition to M, C meets O_1 on each of the arcs A_1B_1, B_1P, $P'B'_1$, $B'_1A'_1$, which contradicts the assumption of at most four intersections. If O is contained in $\mathfrak{R}$ it is necessarily a complete geodesic circle by Lemma 3. 7. If it meets C in some point of C other than M, this is a point of tangency. The deformation of O discussed above will take this point outside of C and yield the same contradiction as before. It follows that O is a complete geodesic circle in $\mathfrak{R}$ meeting C only at the vertex M.

The remainder of the proof of this theorem, based on Lemma 3. 2, is identical with that for the case of the plane [6, Theorem 6. 1], and need not be repeated.

THEOREM 11. 2. *Let a simple closed curve C of class C'' in $\mathcal{S}$, not a geodesic circle, be met by a complete geodesic circle O. If, among the arcs into which O divides C, there can be found n arcs $P_{2i-1}P_{2i}$ $(i = 1, 2, \cdots, n)$ interior to O such that the points P_k are in the same cyclic order on C and O, then C has at least $2n$ geodesic vertices.*

It is possible to restrict attention entirely to the case when none of the P_k are points of tangency, since in any case this can be arranged by a slight deformation of O. It should be noted that the theorem does not require that the arcs selected be all the arcs of C interior to O.

The proof of this theorem is identical with the proof of the corresponding result in the plane [6, Theorem 7.1] except that in the present case we have assumed outright the existence of the n interior arcs. For the present proof Lemma 3.2 and Lemma 5.2 replace Lemma 3.1 and Corollary 4.1.1 respectively of the former paper. The proof will not be duplicated here. The following immediate corollary is easier to visualize, though somewhat less general.

COROLLARY 11.2.1. *If a simple closed curve C of class C'' in $\mathcal{S}$, not a geodesic circle, intersects a complete geodesic circle O in just $2n$ points, and if these intersections have the same cyclic order on C and O, then C has at least $2n$ geodesic vertices.*

In comparing this last theorem with the corresponding results in the plane, at least two questions naturally arise, as follows.

(a) Is Theorem 11.2 still true if O is taken as any geodesic circle, rather than a complete geodesic circle?

(b) The proof of Theorem 11.2 is based essentially on Lemma 5.2 which assures us of the existence of a certain type of vertex on an arc AB tangent to a complete geodesic circle O. Is this lemma itself true if O is any geodesic circle rather than a complete geodesic circle?

We shall proceed to show by counterexamples that both questions are to be answered in the negative. In other words, the restrictions indicated in the theorem above are essential. For convenience the examples will be constructed on the Riemann surface of **9**, but as noted there, they can be realized on an ordinary surface.

Consider any ellipse $x^2/a^2 + y^2/b^2 = 1$ whose eccentricity is greater than $1/\sqrt{2}$, so that the osculating circles at the points $(\pm a, 0)$ do not intersect. Let this figure be subjected to a direct circular transformation taking the center of curvature at $(-a, 0)$ into the point at infinity and taking the x-axis into itself. The result is shown in Figure 2. The two solid circles are the transforms of the circles of curvature at $(\pm a, 0)$ and, as the curve is directed, are the two circles of minimum curvature. Let them be denoted by C_1 and C_2 as shown. Consider the curve C obtained by tracing arc ABD, then tracing C_1 n times, then arc DEA, and finally C_2 n times, the directions being as indicated. As in Section **9**, consider this curve on the logarithmic Riemann surface with branch point Q, as shown. Since the total rotation of the vector QP as P traces C is clearly zero, C is a closed curve on this surface. Moreover, it is easily verified that on this surface it is simple, for the various double points in the plane correspond to points

on different sheets of the surface. C is therefore a simple closed curve with exactly four vertices (since that is true of the ellipse) lying in a simply connected region of a surface of constant (zero) curvature. The simply connected region is that part of the surface covering an annular region of the plane about Q.

Consider the dotted circle K in the figure. Since it goes around the branch point, it is an open arc and has points on all sheets of the Riemann surface. The same is true of C_1, and C_1 and K have two intersections on each sheet of the surface. Since curve C contains the part of C_1 on n sheets,

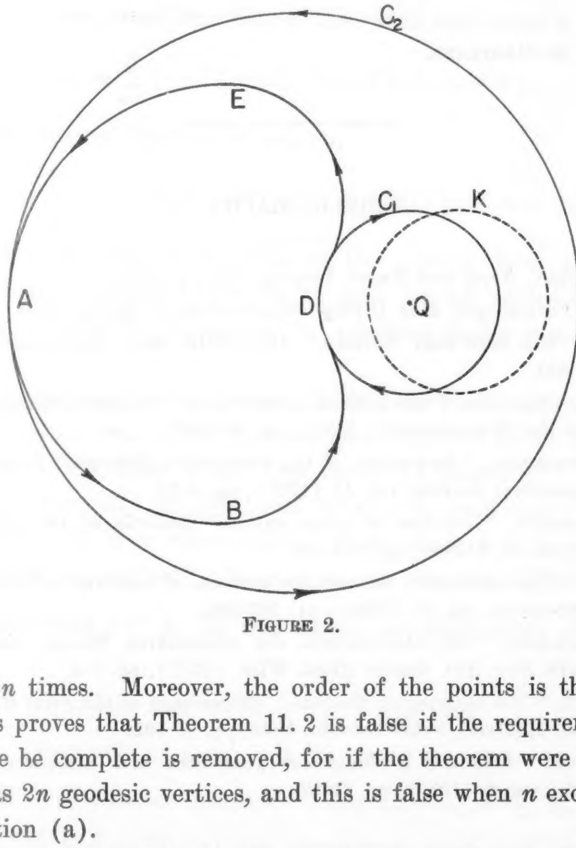


FIGURE 2.

C meets K $2n$ times. Moreover, the order of the points is the same on C and K. This proves that Theorem 11. 2 is false if the requirement that the geodesic circle be complete is removed, for if the theorem were true it would say that C has $2n$ geodesic vertices, and this is false when n exceeds 2. This answers question (a).

To answer question (b) consider the trisectrix whose polar equation is $r = 1 + 2 \operatorname{Cos} \theta$. The smallest circle C which contains this curve has two points of contact with it which divides the trisectrix into two arcs. Let AB be the one of these arcs containing the double point, and let it be directed so that the loop is positively traced. The only vertex on this arc is therefore a maximum. As in the last example, let this be considered on a logarithmic

Riemann surface with branch point at a point O interior to the loop, and therefore also to C. On this surface C is an open arc with points on all sheets. Moreover, on this surface arc AB is simple since the double point corresponds to points on different sheets of the surface. On the surface, therefore, AB is a simple arc tangent to a geodesic circle in the same sense at A and B and lying to the left of this directed circle, yet it has no minimum of the geodesic curvature as there would have to be if Lemma 5.2 were true in this case. Indeed the example is a counterexample even for the simpler Lemma 5.1. Question (b) above must therefore also be answered in the negative.

University of Maryland.

BIBLIOGRAPHY.

1. W. Blaschke, *Kreis and Kugel*, Leipzig, 1916, p. 161.

2. ———, *Vorlesungen über Differentialgeometrie I*, Berlin, 1930.

3. G. Bol, "Ein Satz über Eilinien," *Abh. Math. Sem. Hansischen Univ.*, vol. 13 (1940), pp. 319-320.

4. D. Fog, "Über den Vierscheitelsatz und seine Verallgemeinerungen," *Sitz. der Berlin Akademic der Wissenschaft* (1933), pp. 251-254.

5. W. C. Graustein, "Extensions of the four-vertex theorem," *Transactions of the American Mathematical Society*, vol. 41 (1937), pp. 9-23.

6. S. B. Jackson, "Vertices of plane curves," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 564-578.

7. ———, "The four-vertex theorem for surfaces of constant curvature," *American Journal of Mathematics*, vol. 67 (1945), pp. 563-582.

8. H. Mohrmann, "Die Minimalzahl der stationären Ebenen einer räumlichen Ovals," *Sitz. Math.-Nat. Abt. Bayer. Akad. Wiss.* (1917), pp. 1-4.

9. P. Scherk, "The four-vertex theorem," *Proceedings of the First Canadian Mathematical Congress*, Montreal 1945, Toronto, 1946, pp. 97-102.

10. H. Hopf, "Über die Drehung der Tangenten und Sehnen ebener Kurven," *Compositio Math.*, vol. 2 (1935), pp. 50-62.

THE GENERAL TERM OF THE GENERALIZED SCHLÖMILCH SERIES.*

By J. ERNEST WILKINS, JR.

1. Introduction. By a generalized Schlömilch series is meant a series of the form

$$\tfrac{1}{2}a_0/\Gamma(\nu+1)+\sum_{m=1}^{\infty}(\tfrac{1}{2}mx)^{-\nu}\{a_mJ_\nu(mx)+b_mH_\nu(mx)\},$$

in which $J_\nu(u)$ is the Bessel function of first kind and order ν,

$$J_\nu(u)=\sum_{n=0}^{\infty}(-1)^n(\tfrac{1}{2}u)^{\nu+2n}/n!\Gamma(\nu+n+1),$$

and $H_\nu(u)$ is the Struve function of order ν,

$$H_\nu(u)=\sum_{n=0}^{\infty}(-1)^n(\tfrac{1}{2}u)^{\nu+2n+1}/\Gamma(n+3/2)\Gamma(\nu+n+3/2).$$

Watson [2; 645] has shown that if the general term of the above generalized Schlömilch series converges to zero for all values of x in any interval, then $a_m=o(m^{\nu+\frac{1}{2}})$, $b_m=o(m^{\nu+\frac{1}{2}})$, provided that ν is a real number less than $\frac{1}{2}$.

By analogy with the Cantor lemma [1; 84] we would expect to be able to prove this assertion even if the interval were replaced by an arbitrary set E of positive measure. This is our first result (Theorem 1). By making use of the explicit formulas for $J^{\frac{1}{2}}(u)$ and $H^{\frac{1}{2}}(u)$ it is next seen (Theorem 2) that this result is still true when $\nu=\frac{1}{2}$. Even more is true when $\nu>\frac{1}{2}$; we shall prove that $a_m=o(m^{\nu+\frac{1}{2}})$, $b_m=o(m)$, in this case (Theorem 3).

2. The case when $\nu<\frac{1}{2}$. In this section we shall prove the following theorem.

THEOREM 1. *Suppose that* $-\infty<\nu<\frac{1}{2}$ *and that*

$$(\tfrac{1}{2}mx)^{-\nu}\{a_mJ_\nu(mx)+b_mH_\nu(mx)\}=o(1)$$

for all x in a set E of positive measure. Then $a_m=o(m^{\nu+\frac{1}{2}})$, $b_m=o(m^{\nu+\frac{1}{2}})$.

Since $u^{-\nu}J_\nu(u)$ and $u^{-\nu}H_\nu(u)$ are respectively even and odd functions of u, we can suppose that in the proof of Theorem 1 (and the subsequent theorem also) E is contained in an interval (α,β) such that $0<\alpha<\beta$.

If Theorem 1 were false there would be an infinite subset M of the set of positive integers and a positive quantity η such that $|a_m|^2+|b_m|^2>\eta^2m^{2\nu+1}$

---

* Received August 24, 1948.

if m is in M. Since $0 < \alpha \leqq x \leqq \beta$ if x is in E, it follows that if $f_m(x) = (\pi m x)^{\frac{1}{2}}\{A_m J_\nu(mx) + B_m H_\nu(mx)\}$, where $A_m = a_m/(|a_m|^2 + |b_m|^2)^{\frac{1}{2}}$, $B_m = b_m/(|a_m|^2 + |b_m|^2)^{\frac{1}{2}}$, then $f_m(x) = o(1)$ for all x in E as m approaches ∞ in M. It is known [2; 199] that

$$\begin{aligned} J_\nu(u) &= (2/\pi u)^{\frac{1}{2}}[\cos(u - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + O(u^{-1})], \\ Y_\nu(u) &= (2/\pi u)^{\frac{1}{2}}[\sin(u - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi) + O(u^{-1})], \end{aligned} \tag{2.1}$$

as u approaches ∞, and that [2; 333]

$$H_\nu(u) = Y_\nu(u) + (\tfrac{1}{2}u)^{\nu-1}/\pi^{\frac{1}{2}}\Gamma(\nu + \tfrac{1}{2}) + O(u^{\nu-3}) \tag{2.2}$$

as u approaches ∞. It follows from these relations and the fact that E is bounded away from zero that $J_\nu(mx) = O(m^{-\frac{1}{2}})$ and $H_\nu(mx) = O(m^{-\frac{1}{2}})$ when $\nu < \frac{1}{2}$, and the constant implied by the symbols O can be chosen independent of x. It follows that $f_m(x)$ is bounded, and hence that

$$\int_E |f_m(x)|^2 dx = o(1) \tag{2.3}$$

as m approaches ∞ in M. If we use equations (2.1) and (2.2) we find that

$$\begin{aligned} |f_m(x)|^2 = 1 &+ (|A_m|^2 - |B_m|^2)\sin(2mx - \nu\pi) \\ &- 2Re(A_m B_m^*)\cos(2mx - \nu\pi) + o(1), \end{aligned} \tag{2.4}$$

in which B_m^* is the complex conjugate of B_m, $Re(u)$ is the real part of u, and the term $o(1)$ is uniform in x on E. Since $|A_m| \leqq 1$ and $|B_m| \leqq 1$, it follows upon substituting equation (2.4) into equation (2.3) and using the Riemann-Lebesgue lemma that the measure of E is zero, and this contradicts the hypothesis of the theorem. This contradiction completes the proof of Theorem 1.

3. The case when $\nu = \frac{1}{2}$. In this section we shall prove the following theorem.

THEOREM 2. *The conclusion of Theorem* 1 *follows from its hypothesis if* $\nu = \frac{1}{2}$.

It is known [2; 54, 333] that

$$J_{\frac{1}{2}}(mx) = (2/\pi mx)^{\frac{1}{2}}\sin mx, \qquad H_{\frac{1}{2}}(mx) = (2/\pi mx)^{\frac{1}{2}}(1 - \cos mx).$$

The hypothesis of Theorem 1 thus reduces to the assumption that

$$\begin{aligned} &(mx)^{-1}\{a_m \sin mx + b_m(1 - \cos mx)\} \\ &\qquad = 2(mx)^{-1}\sin\tfrac{1}{2}mx\{a_m\cos\tfrac{1}{2}mx + b_m\sin\tfrac{1}{2}mx\} = o(1) \end{aligned}$$

for all x in E.

If Theorem 2 were false there would be an infinite subset M of the set

of positive integers and a positive quantity η such that $|a_m|^2 + |b_m|^2 > \eta^2 m^2$ whenever m is in M. If E is bounded and bounded away from zero as in the previous section, and if A_m and B_m are defined by equations (2. 0), it follows that $f_m(x) = 2(A_m \cos \frac{1}{2}mx + B_m \sin \frac{1}{2}mx)\sin \frac{1}{2}mx = o(1)$ for all x in E as m approaches ∞ in M. Since $f_m(x)$ is plainly bounded, we have (2. 3) as m approaches ∞ in M. It is easy to see that

$$|f_m(x)|^2 = \tfrac{1}{2}(|A_m|^2 + 3|B_m|^2) - 2|B_m|^2 \cos mx + 2Re(A_m B_m^*)\sin mx$$
$$- Re(A_m B_m^*)\sin 2mx - \tfrac{1}{2}(|A_m|^2 - |B_m|^2)\cos 2mx.$$

It then follows from equation (2. 3) and the Riemann-Lebesgue lemma that $|A_m|^2 + 3|B_m|^2 = o(1)$ as m approaches ∞ in M, since E has positive measure. This is impossible, however, since $|A_m|^2 + 3|B_m|^2 \geqq |A_m|^2 + |B_m|^2 = 1$. This contradiction completes the proof of Theorem 2.

4. The case when $\nu > \frac{1}{2}$. In this section we shall prove the following theorem.

THEOREM 3. *If $\nu > \frac{1}{2}$ and the hypothesis of Theorem* 1 *holds, then* $a_m = o(m^{\nu+\frac{1}{2}})$, $b_m = o(m)$.

If Theorem 3 were false there would be an infinite subset M of the set of positive integers and a positive quantity η such that $c_m^2 = |a_m|^2 m^{-2\nu-1} + |b_m|^2 m^{-2} > \eta^2$ if m is in M. Let us define $A_m = a_m/m^{\nu+\frac{1}{2}}c_m$, $B_m = b_m/m^{\nu+\frac{1}{2}}c_m$. Then

$$|A_m| \leqq 1, \qquad |B_m| \leqq m^{\frac{1}{2}-\nu}. \tag{4.0}$$

If we suppose as in the preceding sections that E is bounded and bounded away from zero it then follows that for every value of μ,

$$f_{m\mu}(x) = x^{(1+\mu)/2}(m\pi)^{\frac{1}{2}}\{A_m J_\nu(mx) + B_m H_\nu(mx)\} = o(1)$$

as m approaches ∞ in M. To see that $f_{m\mu}(x)$ is bounded we use the inequalities (4. 0) and notice from equations (2. 1) and (2. 2) that $J_\nu(mx) = O(m^{-\frac{1}{2}})$ and $H_\nu(mx) = O(m^{\nu-1})$ uniformly in x on E when $\nu > \frac{1}{2}$. It follows that

$$\int_E |f_{m\mu}(x)|^2 dx = o(1)$$

as m approaches ∞ in M. If we define $\phi_\nu(mx)$ so that $H_\nu(mx) = Y_\nu(mx) + \phi_\nu(mx)$, then it follows from equation (2. 2) that

$$\phi_\nu(mx) - (\tfrac{1}{2}mx)^{\nu-1}\pi^{-\frac{1}{2}}/\Gamma(\nu + \tfrac{1}{2}) = O(m^{\nu-3}), \tag{4.1}$$

and that $|f_{m\mu}(x)|^2 = \pi m x^{1+\mu}[|A_m J_\nu(mx) + B_m Y_\nu(mx)|^2 - |B_m|^2 \phi_\nu^2(mx) + 2Re\{B_m^* \phi_\nu(mx)[A_m J_\nu(mx) + B_m H_\nu(mx)]\}]$. If we use equations (2. 2) and (4. 1) we find that

$$
\begin{aligned}
| f_{m\mu}(x) |^2 = x^\mu[| A_m |^2 + | B_m |^2 + (| A_m |^2 - | B_m |^2) \sin(2mx - \nu\pi) \\
- | B_m |^2 \{ (mx)^{2\nu-1} u_\nu^{-1} + O(m^{2\nu-3}) \} - 2Re(A_m B_m^*) \cos(2mx - \nu\pi) \\
+ 2Re\{ B_m^* \phi_\nu(mx) [A_m J_\nu(mx) + B_m H_\nu(mx)] \} + o(1)],
\end{aligned}
$$

in which $u_\nu = 2^{2\nu-2}\Gamma^2(\nu + \frac{1}{2})$. Since (4.0) holds, it follows from the Riemann-Lebesgue lemma that the integral of $x^\mu[(| A_m |^2 - | B_m |^2) \sin(2mx - \nu\pi) - 2Re(A_m B_m^*) \cos(2mx - \nu\pi)]$ over E converges to zero as m approaches ∞ in M. Since $B_m^* \phi_\nu(mx) = O(m^{-\frac{1}{2}})$ uniformly on E, it is true that $x^\mu B_m^* \phi_\nu(mx) [A_m J_\nu(mx) + B_m H_\nu(mx)] = O(m^{-1})$ uniformly on E as m approaches ∞ in M, and hence that the integral of this expression converges to zero. Similarly, the term $| B_m |^2 O(m^{2\nu-3}) = O(m^{-2})$ has an integral over E which converges to zero as m approaches ∞ in M. We thus conclude that

$$
\int_E x^\mu (| A_m |^2 + | B_m |^2 - | B_m |^2 m^{2\nu-1} x^{2\nu-1} u_\nu) \, dx = o(1)
$$

as m approaches ∞ in M. Since $| A_m |^2 + | B_m |^2 \leqq 1$ and $| B_m |^2 m^{2\nu-1} \leqq 1$, we may replace M by an infinite subset for which there exist quantities A and B such that

$$
\lim (| A_m |^2 + | B_m |^2) = A, \qquad \lim | B_m |^2 m^{2\nu-1} = B.
$$

Evidently, $0 \leqq A \leqq 1$, $0 \leqq B \leqq 1$, and

$$
\int_E x^\mu (A - B u_\nu x^{2\nu-1}) \, dx = 0.
$$

Since μ is arbitrary we infer that $A - B u_\nu x^{2\nu-1} = 0$ almost everywhere on E and hence that $A = B = 0$ since E has positive measure. It follows, however, from the definitions of A_m and B_m that

$$
| A_m |^2 + | B_m |^2 (1 + m^{2\nu-1}) = 1 + | b_m |^2 / (| a_m |^2 + | b_m |^2 m^{2\nu-1}) \geqq 1,
$$

whence $A + B \geqq 1$ and so it cannot happen that $A = B = 0$. This contradiction completes the proof of the theorem.

BIBLIOGRAPHY.

1. G. H. Hardy and W. W. Rogosinski, *Fourier series*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 38, Cambridge, 1944.
2. G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge, 1945.

ON THE EXTENSION OF THE PARTIAL ORDER OF GROUPS.*

By Ladislas Fuchs.

1. In his paper "Sur l'extension de l'ordre partiel"[1] E. Szpilrajn has proved that every partial order defined on a set has a linear extension. This general result does not necessarily hold for partially ordered groups, since it may obviously happen that the extended order does not satisfy the group axioms. The principal purpose of the present paper is the demonstration of the same theorem on abelian groups. The theorem will be proved to hold only for groups on which an additional condition is satisfied, a condition which requires that only positive elements have positive natural multiples.

2. We recall that an abelian group G, written additively, is said to be a *partially ordered group*[2] if a relation $>$ is defined between some pairs of its elements such that the following postulates hold:

(i) *any two of the three relations $a > b$, $a = b$, $a < b$ are contradictory*;

(ii) *transitivity*: *$a > b$ and $b > c$ imply $a > c$*;

(iii) *homogeneity*: *$a > b$ implies $a + c > b + c$ for every c in G.*

By the laws (ii) and (iii) the relations $a > b$, $c > d$ may be added to get $a + c > b + d$.

If in addition G satisfies the condition:

(iv) $na = a + a + \cdots + a \geqq 0$ *for some positive integer n implies* $a \geqq 0$,

we say the partial order is *normal.*

If G is such that any two elements a, b are *comparable* in the sense that one of the possibilities $a > b$, $a = b$, $a < b$ does hold, we say G is *linearly ordered.* A linearly ordered group always satisfies condition (iv), for if, under the hypothesis $na \geqq 0$, $a \geqq 0$ did not hold, we should then have by linear order $a < 0$ implying $na < 0$ for every positive integer n.

We now prove that a group on which a normal partial order is defined has, with the exception of 0, only elements of infinite order. In fact, the normality states that if $na = 0$, or, what is the same, if $na \geqq 0$ and $-na \geqq 0$, then $a \geqq 0$ as well as $-a \geqq 0$, that is, $a = 0$.

* Received February 3, 1949.

[1] *Fundamenta Mathematicae*, vol. 16 (1930), pp. 386-389.

[2] Our definition is stated in a form which is most convenient for our purpose; cf. C. J. Everett and S. Ulam, "On ordered groups," *Transactions of the American Mathematical Society*, vol. 57 (1945), pp. 208-216.

Suppose that two partial orders P and R are defined on the same group and a relation $a > b$ in P implies $a > b$ in R; then R will be called an *extension* of P. An extension which defines a linear order on G will be called a *linear extension.*

3. We shall now prove the following lemma.

LEMMA. *If P is a normal partial order on the group G and x and y are any two elements non-comparable in P, then there exists an extension R of P such that $x > y$ in R. Moreover if such an extension may be carried out for any two non-comparable elements, then P is necessarily normal.*[3]

For the proof assume that P is a normal partial order on G and the elements x and y are not comparable in P. Let us define a relation R as follows.

We put $a > b$ in R if and only if $a \neq b$ and there are two non-negative integers p, q, not both zero, such that

$$p(a-b) \geqq q(x-y) \text{ in } P. \tag{1}$$

What we have to show is that by this definition R is a partial order and an extension of P, further $x > y$ in R.

First of all, we note that p is never zero, for otherwise we should have $0 \geqq q(x-y)$ in P for a certain positive integer q, whence by (iv) we have $y \geqq x$ in P against hypothesis.

α) We begin with verifying condition (i) for R. It is clearly enough to show that $a > b$ in R and $b > a$ in R are contradictory. For assume $p(a-b) \geqq q(x-y)$ in P, as well as $r(b-a) \geqq s(x-y)$ in P, for some non-negative integers p, q, r, s. By adding r times the first, p times the second inequality, one obtains $pr(a-b) + pr(b-a) \geqq (qr+ps)(x-y)$ in P, that is to say, $0 \geqq (qr+ps)(x-y)$ in P. If $qr+ps$ does not vanish, by normality we are led to $y \geqq x$ in P, a contradiction. On the other hand, if $qr+ps$ is zero, i. e., both q and s vanish, then the inequalities $p(a-b) \geqq 0$ in P and $r(b-a) \geqq 0$ in P imply by normality $a \geqq b$ in P and $b \geqq a$ in P, i. e., $a = b$ which is absurd.

β) We proceed now to the proof of the transitivity of R. Assume that $a > b$ in R and $b > c$ in R, that is, for some non-negative integers p, q, r, s, $p(a-b) \geqq q(x-y)$ in P and $r(b-c) \geqq s(x-y)$ in P. By adding as in α) one gets $pr(a-b) + pr(b-c) = pr(a-c) \geqq (qr+ps)(x-y)$ in P. Here pr is not zero and $a = c$ is by α) impossible, so that, by definition, $a > c$ in R, which establishes the transitivity of R.

---

[3] Our proof is in some respect similar to Szpilrajn's proof.

γ) In order to prove the homogeneity for R, take into account that (1) includes only the difference $a - b$ which is equal to $(a + c) - (b + c)$ for every c in G and $a + c = b + c$ is impossible if a and b are different.

δ) R is an extension of P, for if $a > b$ in P, then $a - b > 0$ in P, consequently, for $p = 1$, $q = 0$, condition (1) is satisfied, hence $a > b$ in R.

ϵ) One sees at once that $x > y$ in R. In fact, for $x = a$, $y = b$ and $p = q = 1$, the relation (1) takes the form $x - y \geqq x - y$ in P.

ξ) To complete the proof of the lemma, let us now assume that there is an element g in G such that $ng \geqq 0$ in P without $g \geqq 0$ in P. Then g and 0 are not comparable in P and we may infer, from the above discussions, that there exists an extension of P in which $g < 0$. This is however absurd, since this would imply $ng < 0$ in R, contrary to the hypothesis $ng \geqq 0$ in P.

4. We may prove even the normality for the partial order R defined in **3**. Indeed, supposing $na \geqq 0$ in R for some positive integer n, i. e., $p(na) = (pn)a \geqq q(x - y)$ in P, we are led at once to the result $a \geqq 0$ in R.

What has been proved shows that the extended partial order R may again be extended to another partial order S, in which two prescribed elements non-comparable in R become comparable, etc.

If, in general, $P_1, P_2, \cdots, P_\tau, \cdots$ is a well-ordered chain of partial orders such that each of them is some extension of the preceding ones, then the *union* of the chain may be defined to be a partial order P such that $a > b$ in P if and only if $a > b$ in P_τ holds for one and hence for all subsequent subscripts τ. There is no difficulty in establishing that P is normal if all P_τ are normal.

Hence, as a simple consequence of Zorn's lemma [4] we get that in the set of all normal partial orders defined on the group G which are extensions of P, there are maximal orders M, that is, orders which have no proper extension. By our Lemma this can happen only in case any two elements are comparable in M, that is to say, M is a linear order. Thus we have, immediately, the result:

THEOREM 1. *For every normal partial order P defined on G and every two elements x, y non-comparable in P, there exists a linear extension L_{xy} with the property that $x > y$ in L_{xy}.*

---

[4] M. Zorn, "A remark on method in transfinite algebra," *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 667-670.

5. Let $\mathfrak{S} = \{P_1, P_2, \cdots, P_\tau, \cdots\}$ be any set of partial orders, each defined on the same group G. We define a new partial order P on G as follows. For any two elements a, b we put $a > b$ in P if and only if $a > b$ in every partial order P_τ in the set $\mathfrak{S}$. It is readily seen that P is again a partial order, moreover, P is normal if all partial orders of $\mathfrak{S}$ are normal. The partial order P is said to be the *product* of the P_τ or to be *realized* by the set $\mathfrak{S}$ of partial orders, written $P = \Pi P_\tau$.

THEOREM 2. *A partial order P defined on a group G may be realized by a certain set of linear orders if and only if P is normal.*

The necessity is obvious, since a linear order, and hence every product of linear orders, is normal. On the other hand, if P is not itself linear, then take to any pair of elements x, y non-comparable in P the corresponding linear extensions L_{xy} and L_{yx} described in Theorem 1. It is easily seen that these linear orders realize P.

If by the dimension[5] of a partial order P we mean the least cardinal number $\mathfrak{w}$ such that P may be realized by $\mathfrak{w}$ linear orders, then we can reformulate Theorem 2:

THEOREM 3. *A partial order P defined on a group has a dimension if and only if P is normal.*

This theorem states, for example, that each commutative group which is lattice-ordered in the sense of G. Birkhoff,[6] has a well-defined dimension.

6. When we start with an abelian group on which the relation $>$ is defined for no pair of elements, then, by applying Theorem 2 we come to a theorem due to F. Levi.[7]

THEOREM 4. *In a commutative group a linear order may be defined, if all of its elements, except 0, are of infinite order.*

Indeed, a group in which no partial order is defined is normal if and only if it has no element of finite order other than 0.

BUDAPEST, HUNGARY.

---

[5] This definition is due to Ben Dushnik and E. W. Miller, "Partially ordered sets," *American Journal of Mathematics*, vol. 63 (1941), pp. 600-610.

[6] G. Birkhoff, "Lattice-ordered groups," *Annals of Mathematics* (2), vol. 43 (1942), pp. 298-331. Lemma 3 of §9 states that a lattice-ordered abelian group is always normal.

[7] F. Levi, "Arithmetische Gesetze im Gebiete diskreter Gruppen," *Rendiconti Palermo*, vol. 35 (1913), pp. 225-236.

ON THE CONSTRUCTION OF PARTIALLY ORDERED SYSTEMS WITH A GIVEN GROUP OF AUTOMORPHISMS.*

By Robert Frucht.

In a recent paper [1] G. Birkhoff showed that when an abstract group of g elements is given (g being a finite number or an infinite cardinal number), there exists always a partially ordered system with $g^2 + g$ elements whose group of automorphisms is simply isomorphic to the given group. I shall prove the following result for the case of a finite g: *There can be found a partially ordered system with only* $(n+2)g$ *elements whose group of automorphisms is simply isomorphic to a given abstract group of finite order* g, *when this group can be generated by* n *of its elements*; and if $n > 2$, still fewer elements will be needed.

Proof. Let a_1 be the identity of the given group $\mathfrak{G}$, and $a_2, a_3, \cdots, a_{n+1}$ the n generating elements; $a_{n+2}, a_{n+3}, \cdots, a_g$ be the other elements of $\mathfrak{G}$. We proceed now to construct a partially ordered system with the following $(n+2)g$ elements: the g maximal elements (a_1), (a_2), $(a_3), \cdots, (a_g)$ corresponding to the g elements of the group $\mathfrak{G}$, and the $(n+1)g$ other elements (a_ρ, a_σ) which correspond to ordered pairs of elements of $\mathfrak{G}$; here the "first component" a_ρ (where $\rho = 1, 2, \cdots, g$) corresponds to any element of $\mathfrak{G}$, but the "second component" a_σ (where $\sigma = 1, 2, \cdots, n+1$) corresponds either to the unit a_1 or to any *generating* element $a_2, a_3, \cdots, a_{n+1}$ of $\mathfrak{G}$. (This limitation of the second component to the generating elements of $\mathfrak{G}$ just represents our modification of Birkhoff's original method by which the number of elements is reduced from $(g+1)g$ to $(n+2)g$.) Between these $(n+2)g$ elements the following covering relations are defined: [2]

---

* Received May 11, 1949.

[1] Garrett Birkhoff, "Sobre los grupos de automorfismos," *Revista de la Unión Matemática Argentina*, vol. 11 (1946), pp. 155-157; see also: R. Frucht, "Sobre la construcción de sistemas parcialmente ordenados con grupo de automorfismos dado," *Revista de la Unión Matemática Argentina*, vol. 13 (1948), pp. 12-18.

[2] It is not necessary to postulate (b) also for $\tau = 1$, as a_1 is the unit of $\mathfrak{G}$, and therefore $a_1 a_\sigma = a_\sigma$, so that (b) for $\tau = 1$ reads: $(a_\sigma) > (a_\sigma, a_1)$; but this relation is already contained in (a).

$$\begin{cases} \text{(a)} \quad (a_\rho) > (a_\rho, a_1) > (a_\rho, a_2) > \cdots > (a_\rho, a_{n+1}) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (\text{for } \rho = 1, 2, \cdots, g); \\ \text{(b)} \quad (a_\sigma) > (a_\tau a_\sigma, a_\tau) \qquad (\text{for } \sigma = 1, 2, \cdots, g; \tau = 2, 3, \cdots, n+1). \end{cases}$$

It is obvious that the system thus defined is a finite partially ordered system (in the usual[3] terminology), in which no element is isolated. We are now going to prove that this system has a group of automorphisms simply isomorphic to $\mathfrak{G}$.

Let λ be any number from $1, 2, \cdots, g$, and consider the following mapping of the partially ordered system defined by (a) and (b) into itself:

$$\text{(c)} \qquad \Phi_\lambda \begin{cases} (a_\rho) \to (a_\rho a_\lambda) \\ (a_\sigma, a_\tau) \to (a_\sigma a_\lambda, a_\tau). \end{cases}$$

This mapping Φ_λ is clearly order-preserving; e. g. Φ_λ carries (a_σ) into $(a_\sigma a_\lambda)$ and $(a_\tau a_\sigma, a_\tau)$ into $(a_\tau a_\sigma a_\lambda, a_\tau)$, whence the relations (b) are changed into the (likewise correct) relations: $(a_\sigma a_\lambda) > (a_\tau a_\sigma a_\lambda, a_\tau)$.

Now let λ run through $1, 2, \cdots, g$; then the formulae (c) will give g distinct automorphisms $\Phi_1, \Phi_2, \cdots, \Phi_g$ of the partially ordered system, and from the first line of (c) it is obvious that these mappings constitute a group simply isomorphic to $\mathfrak{G}$.

It remains to be shown that there are no other automorphisms of the system besides those given by (c) for $\lambda = 1, 2, \cdots, g$. Let Φ be any order-preserving mapping of the system into itself; we are to prove that Φ coincides with one of the g mappings Φ_λ. According to (a) and (b), the g elements $(a_1), (a_2), \cdots, (a_g)$ are the only maximal elements of the system; hence Φ can merely permute them among themselves. Let (a_μ) be that maximal element into which (a_1) is carried by Φ; and let Φ_μ^{-1} be the automorphism inverse to Φ_μ. Since Φ_μ changes (a_1) into $(a_1 a_\mu) = (a_\mu)$, Φ_μ^{-1} carries (a_μ) into (a_1); hence the product $\Phi\Phi_\mu^{-1}$ (i. e. the mapping Φ followed by the mapping Φ_μ^{-1}) leaves the element (a_1) fixed. We will show that this mapping leaves unchanged all the elements of the partially ordered system.

Let us begin with the maximal chain (of length $n+1$) that starts with (a_1): $(a_1) > (a_1, a_1) > (a_1, a_2) > (a_1, a_3) > \cdots > (a_1, a_{n+1})$. According to (b), (a_1) covers also the elements $(a_2, a_2), (a_3, a_3), \cdots, (a_{n+1}, a_{n+1})$, but since the second component of each of these elements is different from a_1, there is no other maximal chain beginning with (a_1); and (a_1, a_1), (a_1, a_2),

[3] Garrett Birkhoff, *Lattice Theory*, first ed., p. 5.

$\cdots, (a_1, a_{n+1})$ are left fixed by $\Phi\Phi_\mu^{-1}$, as they constitute the only maximal chain starting with (a_1).

Consider now the elements (a_τ, a_τ), with $\tau = 2, 3, \cdots, n+1$, which are covered by (a_1); they could only be interchanged by $\Phi\Phi_\mu^{-1}$; but the largest chain starting with (a_τ, a_τ) is

$$(a_\tau, a_\tau) > (a_\tau, a_{\tau+1}) > (a_\tau, a_{\tau+2}) > \cdots > (a_\tau, a_{n+1})$$

and hence of length $n + 1 - \tau$, and since this length is different for any two values of τ, no such interchange of the elements (a_τ, a_τ) is possible. It follows that also the elements $(a_2, a_2), (a_3, a_3), \cdots, (a_{n+1}, a_{n+1})$ are left fixed by $\Phi\Phi_\mu^{-1}$.

But since any one of these elements belongs only to one maximal chain (of length $n+1$), viz. (a_τ, a_τ) to the maximal chain:

$$(a_\tau) > (a_\tau, a_1) > (a_\tau, a_2) > \cdots > (a_\tau, a_\tau) > \cdots > (a_\tau, a_{n+1}),$$

also all the other elements of all these maximal chains are left fixed by $\Phi\Phi_\mu^{-1}$.

Thus we have already recognized as fixed the elements (a_σ) and (a_σ, a_τ) for $\sigma = 1, 2, \cdots, n+1$; $\tau = 1, 2, \cdots, n+1$. It remains to be proved that the same is true also for $\sigma = n+2, n+3, \cdots, g$. This may be accomplished by the following reasoning:

According to (b) any element (a_σ)—already recognized as fixed if $\sigma = 1, 2, \cdots, n+1$—covers the elements

$$(a_2 a_\sigma, a_2), (a_3 a_\sigma, a_3), \cdots, (a_{n+1} a_\sigma, a_{n+1}),$$

which must be left unchanged by $\Phi\Phi_\mu^{-1}$ too (due to the different lengths of the longest chain starting with each of them); and with them also the other elements of the maximal chains containing them must remain fixed, i. e. all the elements $(a_\rho a_\sigma)$ and $(a_\rho a_\sigma, a_\tau)$, where $a_\rho a_\sigma$ is a product of any two *generating* elements of the group $\mathfrak{G}$.

Repeating this reasoning for the elements covered by $(a_\rho a_\sigma)$ and their maximal chains, we are led to the conclusion that also the elements $(a_\pi a_\rho a_\sigma)$ and $(a_\pi a_\rho a_\sigma, a_\tau)$ are left unchanged by $\Phi\Phi_\mu^{-1}$, where $a_\pi a_\rho a_\sigma$ is any product of 3 generating elements of $\mathfrak{G}$; and continuing the proof in the same way (or shortening it by complete induction as to the number of generating elements appearing as factors), we will easily recognize as fixed *all* the elements (a_σ) and (a_σ, a_τ), since any a_σ is the product of a finite number of generating elements of $\mathfrak{G}$.

The proof is now readily completed as follows: We have already shown that $\Phi\Phi_\mu^{-1}$ leaves all the elements of our system unchanged; hence $\Phi\Phi_\mu^{-1} = \Phi_1$ (identity), and $\Phi = \Phi_1\Phi_\mu = \Phi_\mu$. Thus we have proved that any automorphism of the system coincides with one of the mappings Φ_μ given by (c), and we already know that these g mappings constitute a group simply isomorphic to the given abstract group $\mathfrak{G}$.

Finally we shall prove that if $n > 2$ the system having the covering relations (a) *and* (b) *defined above may be modified so that less than $(n+2)g$ elements are needed.* To prove this it will suffice to notice that if $n > 2$ it is possible to drop all the elements (a_ρ, a_σ) where σ is greater than a number m depending on n and defined as the least integer satisfying the inequality: $m \geqq \frac{1}{2}[1 + (8n+1)^{\frac{1}{2}}]$. One must only limit the covering relations (a) to the retained elements, and replace those of the relations (b) which refer to dropped elements, by others of the type:

$$\text{(b}'\text{)} \qquad (a_\rho, a_\kappa) > (a_\sigma a_\rho, a_\lambda),$$

where σ runs through $m+1, m+2, \cdots, n+1$, and where for any two values of σ two distinct combinations $\kappa < \lambda$ (from the numbers $1, 2, \cdots, m$) must be chosen. In order to avoid the formation of new maximum chains it will be convenient to choose always $\kappa \leqq \lambda - 2$. As there are $\frac{1}{2}(m-1)(m-2)$ combinations satisfying this condition, and $n - m + 1$ dropped suffixes, m can be chosen as the least positive integer satisfying the inequality:

$$\tfrac{1}{2}(m-1)(m-2) \geqq n - m + 1,$$

whence the value of m given above follows.

The proof that also the modified system of $(m+1)g$ elements has a group of automorphisms simply isomorphic to the given group $\mathfrak{G}$ may easily be supplied by the reader along the same lines of the proof given above for the system defined by (a) and (b).

Addendum.

Of course it will be convenient to take n always as small as possible, e. g. $n = 1$ for cyclic groups, $n = 2$ for dihedral groups, etc. In any group of finite order g however the number n of generators can always be chosen so small that $n \leqq \log g/\log 2$; hence we have the following

COROLLARY. *To obtain a partially ordered system whose group is simply isomorphic to a given group of finite order g, at most $[(2 + \log g/\log 2)g]$ elements are needed.*[4]

Here $[\ldots]$ stands as usual for "greatest integer $\leqq$."

It may be noted that when the factorization of g into prime powers is known: $g = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_k^{\alpha_k}$, it can be shown that

$$n \leqq \alpha_1 + \alpha_2 + \alpha_3 + \cdots + \alpha_k,$$

whence another corollary similar to the foregoing can be concluded.

TECHNICAL UNIVERSITY F. SANTA MARIA,
VALPARAISO, CHILE, SOUTH AMERICA.

---

[4] Combining the two inequalities: $m < \frac{1}{2}\{3 + (8n + 1)^{\frac{1}{2}}\}$ and $n \leqq \log g/\log 2$, it follows that in this corollary $[(2 + \log g/\log 2)g]$ can be replaced by $[(5 + \gamma)g/2]$, where $\gamma = (1 + 8 \log g/\log 2)^{\frac{1}{2}}$.

ON THE BEHAVIOUR OF FOURIER TRANSFORMS AT INFINITY AND ON QUASI-ANALYTIC CLASSES OF FUNCTIONS.* [1]

By I. I. Hirschman, Jr.

1. Introduction. Let $f(x) \varepsilon L_2(-\infty, \infty)$ and $\phi(t) \varepsilon L_2(-\infty, \infty)$ be corresponding Fourier transforms,

$$f(x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{ixt} \phi(t)\, dt \qquad (\mathrm{M}_2)$$

(1)

$$\phi(t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-ixt} f(x)\, dx \qquad (\mathrm{M}_2)$$

The symbol (M_2) indicates that the integrals in equations (1) are to be taken in the sense (2) l. i. m. $T \to \infty \int_{-T}^{T}$. A general principle, due to N. Wiener, states that $f(x)$ and $\phi(t)$ cannot both be too small at infinity unless they are both zero almost everywhere. This principle was first realized as a theorem by G. H. Hardy. A short time later further theorems were obtained by G. W. Morgan. They proved somewhat more precise results which imply that if

$$|\phi(t)| = O(e^{-|t|^p}) \ (t \to \pm \infty); \qquad |f(x)| = O(e^{-|x|^q}) \ (x \to \pm \infty),$$

where p and q are positive and such that $p^{-1} + q^{-1} < 1$, then $f(x)$ and $\phi(t)$ are zero almost everywhere.

We shall consider the case where $\phi(t)$ approaches zero very slowly while $f(x)$ approaches zero very rapidly, a case which has applications to the theory of quasi-analytic functions. It is interesting to give a typical though very special case of our results, the general statement of which we postpone until Section 2. Let

$$\phi(t) e^{|t|\theta(t)} \varepsilon L_2(-\infty, \infty), \tag{2}$$

where $\theta(t) = (\pi/2)[\log(|t| + e)]^{-1}$. If $f(x)$ is the L_2 Fourier transform of $\phi(t)$ and if

$$f(x) = O(\exp\{-\exp\exp(1+\epsilon)|x|\}), \quad (x \to +\infty \text{ or } x \to -\infty), \tag{3}$$

---

* Received March 22, 1949.

[1] Research supported in part by the Office of Naval Research.

for any $\epsilon > 0$, then $f(x) = 0$ almost everywhere. On the other hand, there exists a function $f(x) \not\equiv 0$, the L_2 transform of a function $\phi(t)$ satisfying (2), such that equation (3) holds for every $\epsilon < 0$. We may equally well treat one-sided conditions. Let $\phi(t)e^{|t|\theta(t)} \varepsilon L_2(-\infty, \infty)$, where

$$\theta(t) = \pi[\log(t+e)]^{-1} \qquad t > 0, \qquad \theta(t) = 0 \qquad t \leqq 0.$$

If $f(x)$ is the L_2 transform of $\phi(t)$ and if equation (3) holds for any $\epsilon > 0$, then $f(x)$ is zero almost everywhere.

Let $f(x)$ be an infinitely differentiable function defined for $-\infty < x < \infty$ such that

(4) $$|f^{(n)}(x)| \leqq A(2/\pi)^n n! \qquad (-\infty < x < \infty; n = 0, 1, \cdots).$$

$f(x)$ is then the restriction to the real axis of a function $f(z)$ analytic and bounded in every strip $|\operatorname{Im} z| \leqq l < \pi/2$. A simple application of the Phragmen-Lindelöf principle shows that if

(5) $$f(x) = O(\exp[-\exp(1+\epsilon)x]) \qquad (x \to +\infty)$$

for any $\epsilon > 0$, then $f(x) \equiv 0$. On the other hand there exists a function $f(x) \not\equiv 0$ satisfying (4) and such that equation (5) holds for every $\epsilon < 0$. Using our results on Fourier transforms we shall obtain similar results for quasi-analytic classes of functions defined by the more general inequalities

$$|f^{(n)}(x)| \leqq Ak^n M_n \qquad (-\infty < x < \infty; n = 0, 1, \cdots).$$

Let us consider a special case closely related to our previous examples. If

(6) $$|f^{(n)}(x)| \leqq An!(\log(n+e))^n(2/\pi)^n \quad (-\infty < x < \infty; n = 0, 1, \cdots),$$

and if

(7) $$f(x) = O(\exp\{-\exp\exp(1+\epsilon)x\}) \qquad (x \to +\infty),$$

for $\epsilon > 0$, then $f(x) \equiv 0$. On the other hand, there exists a function $f(x) \not\equiv 0$, satisfying inequalities (6) and such that (7) holds for every $\epsilon < 0$.

2. The main theorem. We proceed immediately to the demonstration of our principal result.

THEOREM 2. *Let*

1. $\phi(t)e^{|t|\theta(t)} = \psi(t) \varepsilon L_2(-\infty, \infty)$
where $0 \leqq \theta(t) \leqq M$, $-\infty < t < \infty$, *for some* M;

2. $H(r) = \pi^{-1} \int_{-r}^{r} |\, t \,|\, \theta(t)[1 + t^2]^{-1} dt,\ H(r) \to \infty$ *as* $r \to \infty$;

3. $f(x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{ixt} \phi(t)\, dt$ (M_2);

4. $f(x) = O(e^{-L(x)})$ $(x \to +\infty)$,

where for x sufficiently large $L(x)$ is positive and $L(x)/x$ is strictly increasing to $+\infty$. If $\overline{\lim_{r\to\infty}} H[L(r)]r^{-1} > 1$, then $f(x) = 0$ almost everywhere.

We define

$$(1) \qquad F(w) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-ixw} f(x)\, dx \qquad (w = u + iv).$$

If $v > 0$ then the integral (1) converges because of assumption 4. The function $F(w)$ is thus analytic for $v > 0$. Moreover since

$$\lim_{v\to 0+} \| e^{vx} f(x) - f(x) \|_2 = 0,$$

we have, by Parseval's theorem,

$$(2) \qquad \lim_{v\to 0+} \| F(u + iv) - \phi(u) \|_2 = 0.$$

Let S_r denote the strip $0 \leqq \operatorname{Im} w \leqq r$. The function $U(w)$ harmonic in S_r which assumes the boundary values $b_1(u)$ for $\operatorname{Im} w = 0$ and $b_2(u)$ for $\operatorname{Im} w = r$ is given by the formula

$$(3) \qquad U(u + iv) = \int_{-\infty}^{\infty} b_1(u - t) k(v, t, r)\, dt + \int_{-\infty}^{\infty} b_2(u - t) k(r - v, t, r)\, dt,$$

where

$$(4) \qquad k(v, t, r) = (2r)^{-1} \tan(\pi v/2r)[\tan^2(\pi v/2r) + \tanh^2(\pi t/2r)]^{-1} \cosh^{-2}(\pi t/2r).$$

This formula is easily verified by means of the mapping $\zeta = \tanh(\pi w/2r)$ which carries the strip S_r into the half plane $\operatorname{Im} \zeta \geqq 0$.

It is evident that the function $F(w)$ is bounded in every strip $\epsilon \leqq \operatorname{Im} w \leqq r$. Since $\log |\, F(w) \,|$ is a subharmonic function and since it is bounded from above we have, by the principle of majoration, that

$$\log |\, F(iv) \,| \leqq \int_{-\infty}^{\infty} \log |\, F(t + i\epsilon) \,|\, k(v - \epsilon, t, r - \epsilon)\, dt$$

$$\int_{-\infty}^{\infty} \log |\, F(t + ir) \,|\, k(r - v, t, r - \epsilon)\, dt.$$

Let us define $M(r) = \operatorname*{Max}_{-\infty < u < \infty} \log |\, F(u + ir) \,|$. It can be verified that

$$\int_{-\infty}^{\infty} k(r-v, t, r-\epsilon)\, dt = (v-\epsilon)/(r-\epsilon).$$

Thus

$$\varlimsup_{\epsilon\to 0+} \int_{-\infty}^{\infty} \log |F(t+ir)|\, k(r-v, t, r-\epsilon)\, dt \leqq (v/r) M(r).$$

Equation (2) implies that $\lim\limits_{\epsilon\to 0+} \| \log^+ F(t+i\epsilon) - \log^+ \phi(t) \|_2 = 0$, and thus that

$$\lim_{\epsilon\to 0+} \int_{-\infty}^{\infty} \log^+ |F(t+i\epsilon)|\, k(v-\epsilon, t, r-\epsilon)\, dt$$
$$= \int_{-\infty}^{\infty} \log^+ |\phi(t)|\, k(v, t, r)\, dt.$$

Finally by what is essentially Fatou's theorem, see [10; p. 346],

$$\lim_{\epsilon\to 0+} \int_{-\infty}^{\infty} \log^- |F(t+i\epsilon)|\, k(v-\epsilon, t, r-\epsilon)\, dt$$
$$\leqq \int_{-\infty}^{\infty} \log^- |\phi(t)|\, k(v, t, r)\, dt.$$

Combining these three equations we see that for $v > 0$

$$\log |F(iv)| \leqq \int_{-\infty}^{\infty} \log |\phi(t)|\, k(v, t, r)\, dt + (v/r) M(r)$$
$$(5) \qquad \leqq -\int_{-\infty}^{\infty} |\, t \,|\, \theta(t) k(v, t, r)\, dt + \int_{-\infty}^{\infty} \log |\psi(t)|\, k(v, t, r) dt + (v/r) M(r).$$

By $A(r) \sim B(r)$ we mean that $\lim\limits_{r\to\infty} A(r)/B(r) = 1$.

LEMMA 2a.

$$\int_{-\infty}^{\infty} |\, t \,|\, \theta(t) k(v, t, r)\, dt \sim vH(r) \qquad (r \to +\infty).$$

We first assert that if $\epsilon > 0$ *then*

$$(6) \qquad \int_{-\infty}^{\infty} |\, t \,|\, \theta(t) k(v, t, r)\, dt \sim \int_{-\epsilon r}^{\epsilon r} |\, t \,|\, \theta(t) k(v, t, r)\, dt \qquad (r \to +\infty).$$

This is an immediate consequence of the relation

$$(7) \qquad \int_{|t| \geqq r\epsilon} |\, t \,|\, \theta(t) k(v, t, r)\, dt = O(1) \qquad (r \to +\infty),$$

which we will now prove. Recalling the definition of $k(v, t, r)$ and making the change of variable $t = rx$ we have

$$\int_{|t|\geqq r\epsilon} |t|\,\theta(t)k(v,t,r) = (r/2)\tan(\pi v/2r)\int_{|x|\geqq\epsilon}|x|\,\theta(rx)[\tan^2(\pi v/2r)$$
$$+\tanh^2(\pi x/2)]^{-1}\cosh^{-2}(\pi x/2)\,dx \leqq \tfrac{1}{2}Mr\tan(\pi v/2r)\int_{|x|\geqq\epsilon}|x|\sinh^{-2}(\pi x/2)\,dx,$$

where M is the constant of assumption 1. Since $\lim_{r\to+\infty}(r/2)\tan(\pi v/2r) = (\pi v/4)$, equation (7) follows.

If $\epsilon > 0$ then we may show by the same argument that

$$(8)\qquad (v/\pi)\int_{-r}^{r}|t|\,\theta(t)[v^2+t^2]^{-1}dt \sim (v/\pi)\int_{-\epsilon r}^{\epsilon r}|t|\,\theta(t)[v^2+t^2]^{-1}dt.$$

We next assert that given $\delta > 0$ there exists $\epsilon > 0$ and r_0 such that

$$(9)\qquad (1-\delta)(v/\pi)(v^2+t^2)^{-1} \leqq k(v,t,r) \leqq (1+\delta)(v/\pi)(v^2+t^2)^{-1} \qquad (|t| \leqq \epsilon r,\ r \geqq r_0).$$

This is easily seen. Combining equations (6), (8), and (9) we have proved that

$$\int_{-\infty}^{\infty}|t|\,\theta(t)k(v,t,r)\,dt \sim (v/\pi)\int_{-r}^{r}|t|\,\theta(t)[v^2+t^2]^{-1}dt.$$

It remains only to show that

$$(v/\pi)\int_{-r}^{r}|t|\,\theta(t)[v^2+t^2]^{-1}dt \sim vH(r).$$

This follows from the fact that

$$vH(r) - (v/\pi\int_{-r}^{r}|t|\,\theta(t)[v\ \ +t^2]^{-1}dt$$
$$= (v/\pi)(v^2-1)\int_{-r}^{r}|t|\,\theta(t)[v^2+t^2]^{-1}[1+t^2]^{-1}dt = O(1) \qquad (r\to+\infty).$$

LEMMA 2b.

$$\overline{\lim_{r\to+\infty}}\int_{-\infty}^{\infty}k(v,t,r)\log|\psi(t)|\,dt < \infty.$$

We have

$$\int_{-\infty}^{\infty}k(v,t,r)\log|\psi(t)|\,dt \leqq \int_{-\infty}^{\infty}k(v,t,r)\log^+|\psi(t)|\,dt.$$

Since $\psi(t)\ \varepsilon\ L_2$, $\log^+|\psi(t)|\ \varepsilon\ L_2$, and hence

$$\overline{\lim_{r\to+\infty}}\int_{-\infty}^{\infty}k(v,t,r)\log|\psi(t)|\,dt \leqq (v/\pi)\int_{-\infty}^{\infty}\log^+|\psi(t)|\,[v^2+t^2]^{-1}dt.$$

For x and r sufficiently large, $x > x_0$, $r > r_0$, there is because of assumption 4 a unique solution $x = \Lambda(r)$ of the equation $xr = L(x)$.

LEMMA 2c.

$$M(r) \leqq \sim (r+1)\Lambda(r+1) \qquad (r \to +\infty).$$

We have

$$\log | F(u+ir) | \leqq \log[(\int_{-\infty}^{x_0} + \int_{x_0}^{\Lambda(r+1)} + \int_{\Lambda(r+1)}^{\infty}) e^{rx} | f(x) | \, dx]$$

$$\leqq \log[I_1(r) + I_2(r) + I_3(r)].$$

Now

$$I_1(r) = O(e^{x_0 r}) \qquad (r \to +\infty).$$

$$I_2(r) \leqq \int_{x_0}^{\Lambda(r+1)} e^{-x} e^{(r+1)x} dx \leqq e^{(r+1)\Lambda(r+1)} \qquad (r \geqq r_0).$$

$$I_3(r) \leqq \int_{\Lambda(r+1)}^{\infty} e^{-x} e^{(r+1)x} e^{-L(x)} dx = o(1) \qquad (r \to +\infty).$$

From these estimates our lemma follows immediately.

LEMMA 2d. *If* $0 < a$, $1 < b$, *then*

$$H(a) \geqq H(ab) - (2M/\pi) \log b, \tag{10}$$

$$H(r+a) \sim H(r) \qquad (r \to +\infty). \tag{11}$$

We have

$$H(ab) = \pi^{-1} \int_{-a}^{a} | t | \, \theta(t) [+ t^2]^{-1} dt + \pi^{-1} \int_{a \leqq |t| \leqq ab} | t | \, \theta(t) [1 + t^2]^{-1} dt$$

$$\leqq H(a) + (2M/\pi) \log b.$$

After transposition this is inequality (10). Setting $b = (r+a)/r$ we find that

$$H(r) \leqq H(r+a) \leqq H(r) + 2M/\pi) \log[(r+a)/r].$$

Equation (11) follows.

Using equation (5) and Lemmas 2a, 2b, and 2c, we see that, if for some $\epsilon > 0$

$$\lim_{r \to +\infty} r^{-1}[(r+1)\Lambda(r+1)] - (1-\epsilon) H(r) = -\infty,$$

or, what is the same thing, if

$$\overline{\lim_{r \to +\infty}} H(r) r / (r+1) \Lambda(r+1) > 1, \tag{12}$$

then $F(w) \equiv 0$ and hence $f(x) = 0$ almost everywhere. By equation (11) of Lemma 2d we see that

$$H(r)r/(r+1)\Lambda(r+1) \sim H(r+1)/\Lambda(r+1).$$

If we set $(r+1) = L(x)/x$, $\Lambda(r+1) = x$, we find that equation (12) is equivalent to

$$\overline{\lim_{x\to\infty}} H[L(x)/x]/x > 1.$$

By equation (10) of Lemma 2d we have $H[L(x)/x]/x \sim H[L(x)]/x$, and our theorem is proved.

3. The converse.

THEOREM 3. *If*

1. $0 \leqq \theta(t) \leqq M \quad (-\infty < t < \infty) \qquad \theta(t) = \theta(-t)$;
2. $t\theta(t) \,\varepsilon \uparrow \quad (0 \leqq t < \infty)$;
3. $H(r) = \pi^{-1}\int_{-r}^{r} |t|\,\theta(t)[1+t^2]^{-1}dt$, $H(r) \to \infty$ as $r \to +\infty$, *then*

there exists a function $\phi(t)$ not equivalent to zero such that

$$|\phi(t)|\, e^{|t|\theta(t)} \leqq [1+t^2]^{-1} \qquad (-\infty < t < \infty)$$

and such that

$$\textit{if } f(x) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{ixt}\phi(t)\,dt, \textit{ then } f(x) = O(\exp - H^{-1}[(1-\epsilon)x]), \quad (x \to +\infty).$$

for every $\epsilon > 0$.

The equation

$$U(u+iv) = (v/\pi)\int_{-\infty}^{\infty} |t|\,\theta(t)\{[1+t^2]^{-1} - [v^2 + (t-u)^2]^{-1}\}dt$$

defines $U(u+iv)$ as a harmonic function in the half plane $v > 0$. For $v = 0$, U assumes the boundary values $|u|\,\theta(u)$, i.e., $\lim_{v\to 0+} U(u+iv) = -|u|\,\theta(u)$ almost everywhere.

We assert that

$$(1) \qquad U(u+iv) \leqq U(iv) \qquad (-\infty < u < \infty; v > 0).$$

We have

$$\partial U(u+iv)/\partial u = v/\pi \int_{-\infty}^{\infty} |t|\,\theta(t)\{-2(t-u)[v^2+(t-u)^2]^{-2}\}dt$$

$$= v/\pi \int_{0}^{\infty} [|u+t|\,\theta(u+t) - |u-t|\,\theta(u-t)]\{-2t[v^2+t^2]^{-1}\}dt.$$

Using assumptions 1. and 2, we see that $\partial U/\partial u$ is negative for $u > 0$ and positive for $u < 0$, and this implies the validity of equation (1).

We will show that

$$(2)\qquad U(iv)v^{-1} \sim H(v) \qquad (v \to +\infty).$$

Now $U(iv)v^{-1} = \pi^{-1}\int_{-\infty}^{\infty} |\, t \,|\, \theta(t)(v^2-1)(v^2+t^2)^{-1}(1+t^2)^{-1}dt.$ If $\epsilon > 0$, then

$$(3)\qquad U(iv)v^{-1} \sim \pi^{-1}\int_{-\epsilon v}^{\epsilon v} |\, t \,|\, \theta(t)(v^2-1)(v^2+t^2)^{-1}(1+t^2)^{-1}dt.$$

This is proved, as usual, by showing that

$$\pi^{-1}\int_{|t| \geqq \epsilon v} |\, t \,|\, \theta(t)(v^2-1)(v^2+t^2)^{-1}(1+t^2)^{-1}dt = O(1) \qquad (v \to +\infty).$$

Making the change of variable $t = vx$, we have

$$\begin{aligned}\pi^{-1}\int_{|t| \geqq \epsilon v} |\, t \,|\, \theta(t)(v^2-1)(v^2+t^2)^{-1}dt \\ &= \pi^{-1}\int_{|x| \geqq \epsilon} |\, x \,|\, \theta(vx)(v^2-1)(1+x^2)^{-1}[1+v^2x^2]^{-1}dx \\ &\leqq M(\pi\epsilon)^{-1}\int_{|x| \geqq \epsilon}(1+x^2)^{-1}dx.\end{aligned}$$

Equation (3) now follows. Given $\delta > 0$ there exists $\epsilon > 0$ and $v_0 > 0$ such that

$$(1-\delta)(1+t^2)^{-1} \leqq (v^2-1)(v^2+t^2)^{-1}(1+t^2)^{-1} \leqq (1+\delta)(1+t^2)^{-2}$$
$$(|\, t \,| \leqq v\epsilon;\, v \geqq v_0).$$

The proof of equation (2) may now be completed exactly in the manner of Lemma 2a.

Let $V(u+iv)$ be a conjugate harmonic function of $U(u+iv)$. The function $\exp[U(w)+iV(w)]$ is analytic for $v > 0$ and is bounded in every strip $0 < v < r$. In particular its modulus does not exceed 1 in the strip $0 < v < 1$. We set

$$F(w) = (w+i)^{-2}\exp[U(w)+iV(w)],$$

and

$$f(x) = (2\pi)^{-1/2}e^{-\xi x}\int_{-\infty}^{\infty} e^{ixt}F(t+i\xi)\,dt \qquad (-\infty < x < \infty;\, \xi > 0).$$

It is easily verified using contour integration that $f(x)$ is independent of ξ.

By Fatou's theorem, see [1; vol. 2, p. 147), $\lim_{\epsilon\to 0+} F(t+i\xi)$ exists for almost all t. Call this limit $\phi(t)$. Since $|F(t+i\xi)| \leqq (1+t^2)^{-1}$ for $(0<\xi<1)$ we have, by the principle of dominated convergence,

$$f(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ixt}\phi(t)\,dt.$$

Moreover $|\phi(t)| \leqq e^{-|t|\theta(t)}(1+t^2)^{-1}$ almost everywhere.

It remains to prove that

$$f(x) = O(\exp\{-H^{-1}[(1-\epsilon)x]\}) \qquad (x\to+\infty)$$

for every $\epsilon>0$. From the definition of $f(x)$ and from equation (1) we see that

$$|f(x)| \leqq 2\pi)^{-\frac{1}{2}} e^{-\xi x} e^{U(i\xi)} \int_{-\infty}^{\infty} [(\xi+1)^2+t^2]^{-1}dt.$$

By equation (2), if ξ is sufficiently large, $|f(x)| \leqq \exp\{-\xi x + (1+\epsilon)\xi H(\xi)\}$. We now set $\xi = H^{-1}[(x-1)/(1+\epsilon)]$ to obtain, for x sufficiently large, $|f(x)| \leqq \exp\{-H^{-1}[(x-1)/(1+\epsilon)]\}$. Since, when x is large, $x(1-\epsilon) \leqq (x-1)/(1+\epsilon)$, we have $|f(x)| \leqq \exp\{-H^{-1}[(1-\epsilon)x]$, and this completes the proof of Theorem 3.

4. Quasi-analytic functions. Let $C\{M_n\}$ denote the class of functions $f(x)$ defined and infinitely differentiable for $-\infty<x<\infty$ and such that

$$(1) \qquad |f^{(n)}(x)| \leqq Ak^n M_n \qquad (-\infty<x<\infty; n=0,1,\cdots),$$

where A and k are constants which may depend upon $f(x)$. The convex regularization $\{M_n{}^c\}$ of the sequence $\{M_n\}$ is defined by the equations

$$(2) \qquad T(r) = \operatorname*{Max}_{n\geqq 0}(r^n/M_n), \qquad M_n{}^c = \operatorname*{Max}_{r\geqq 0}(r^n/T(r)).$$

See [8]. It may be shown that

$$(3) \qquad M_n{}^c \leqq M_n, \qquad M_n{}^c \leqq (M_{n-j}{}^c)^{k/(j+k)}(M_{n+k}{}^c)^{j/(j+k)}.$$

The class $C\{M_n\}$ is said to be quasi-analytic if a function $f(x) \in C\{M_n\}$ which vanishes with all its derivatives at a point $x=x_0$ is necessarily indentically zero. It is well known, see [8], that a necessary and sufficient condition for a class $C\{M_n\}$ to be quasi-analytic is that the integral

$$\int^{\infty} \log T(r)\,r^{-2}dr$$

diverge.

The class $C\{M_n\}$ is said to contain the class $C\{n!\}$, the analytic class, if every $f(x)$ in $C\{n!\}$ belongs also to $C\{M_n\}$. Using the function $f(x) = (i+x)^{-1}$ we see that a necessary and sufficient condition for $C\{M_n\}$ to contain $C\{n!\}$ is that

$$n! \leqq Bb^n M_n \qquad (n = 0, 1, \cdots), \tag{4}$$

for some positive constants B and b.

Kolmogoroff, [5], has shown that, if $m_n = \sup_{-\infty < x < \infty} |f^{(n)}(x)|$,

$$m_n \leqq (\pi/2)(m_{n-j})^{k/(j+k)}(m_{n+k})^{j/(j+k)}.$$

It follows that if $f(x)$ satisfies inequalities (1) then

$$|f^{(n)}(x)| \leqq (\tfrac{1}{2}\pi A)k^n M_n{}^c \qquad (-\infty < x < \infty;\ n = 0, 1, \cdots). \tag{5}$$

THEOREM 4. *Let*

1. $C\{M_n\} \supset C\{n!\}$;
2. $H(r) = (2/\pi)\int^r \log T(u)u^{-2}du,\ H(r) \to \infty$ *as* $r \to +\infty$;
3. $|f^{(n)}(x)| \leqq Ak^n M_n \qquad (-\infty < x < \infty;\ n = 0, 1, \cdots)$.

If $f(x) = O(e^{-L(x)})$, $(x \to +\infty)$, *where* $L(x)$ *is positive and strictly increasing to infinity, and if* $\overline{\lim}_{x\to+\infty} H[L(x)]/x > k$, *then* $f(x) \equiv 0$.

Conversely, if assumptions 1. *and* 2. *are satisfied, then there exists a function* $f(x)$ *satisfying* 3. *and such that* $f(x) = O(\exp - H^{-1}(k'x))$, $(x \to +\infty)$ *for every* $k' < k$.

We define $F(x) = (\sin x/x)f(x)$. We assert that if $k_1 > k$ then there exists a constant A_1 such that

$$\| F^{(n)}(x) \|_2 \leqq A_1 k_1{}^n M_n{}^c \qquad (n = 0, 1, \cdots). \tag{6}$$

We have $F^{(n)}(x) = \sum_{j=0}^{n}\binom{n}{j} f^{(n-j)}(x)(\sin x/x)^{(j)}$. Using equation (5) we obtain

$$\| F^{(n)}(x) \|_2 \leqq \sum_{j=0}^{n}\binom{n}{j}\tfrac{1}{2}\pi A k^{n-j} M_{n-j}{}^c \,\|(\sin x/x)^{(j)}\|_2.$$

Now

$$(\sin x/x)^{(j)} = \tfrac{1}{2}\int_{-1}^{1} e^{ixt}(it)^j dt, \qquad \|(\sin x/x)^{(j)}\|_2 = (\pi/2)^{1/2}[\int_{-1}^{1} t^{2j}dt]^{1/2}$$
$$\leqq (\pi/2)^{1/2} \qquad (j = 0, 1, \cdots).$$

Also, by equation (3), $M_{n-j}^c \leqq (M_n^c)^{1-(j/n)}(M_0^c)^{j/n}$. Thus

$$\begin{aligned}\| F^{(n)}(x)\|_2 &\leqq (\pi/2)^{3/2}A\sum_{j=0}^{n}\binom{n}{j}k^{n-j}(M_n^c)^{1-(j/n)}(M_0^c)^{j/n}\\ &\leqq (\pi/^{3/2}A[(M_n^c)^{1/n}k+(M_0^c)^{1/n}]^n\\ &\leqq (\pi/2)^{3/2}Ak^nM_n^c[1+(M_0^c/M_n^c)^{1/n}k^{-1}]^n.\end{aligned}$$

Since, from assumption 1., $(M_n^c)^{1/n}\to\infty$ as $n\to\infty$, equation (6) follows.

Because $F(x)\,\varepsilon\,L_2(-\infty,\infty)$, there is a function $\phi(t)\,\varepsilon\,L_2(-\infty,\infty)$ such that

$$F(x) = (2\pi)^{-\frac{1}{2}}\int_{-\infty}^{\infty}e^{ixt}\phi(t)\,dt \qquad (\mathrm{M}_2).$$

Moreover,

$$F^{(n)}(x) = (2\pi)^{-\frac{1}{2}}\int_{-\infty}^{\infty}e^{ixt}(it)^n\phi(t)\,dt \qquad (\mathrm{M}_2).$$

This implies that if $k_2 > k_1$, then

$$\|\sum_{n=0}^{\infty}(t/k_2)^nM_n^{-1}\phi(t)\|_2 \leqq \sum_{n=0}^{\infty}k_2^{-n}M_n^{-1}\,\|t^n\phi(t)\|_2 = \sum_{n=0}^{\infty}k_2^{-n}M_n^{-1}\,\|F^{(n)}(x)\|_2$$

$\leqq A_1\sum_{n=0}^{\infty}(k_1/k_2)^n < \infty$. This in turn implies that $\|T(|t|/k_2)\phi(t)\|_2 < \infty$.

Let us define $\theta(t)$ by the equations

$$\theta(t) = 0, \quad (|t| \leqq 1); \qquad |t|\,\theta(t) = \log^+[T(|t|/k_2)], \quad (|t| > 1).$$

We then have $e^{|t|\theta(t)}\phi(t)\,\varepsilon\,L_2(-\infty,\infty)$. Clearly $\theta(t) \geqq 0$. Assumption 1. implies inequality (4) for suitably chosen B and b, and equation (4) implies that $\theta(t) \leqq M$ for a suitably chosen constant M. Finally $F(x) = O(e^{-L(x)})$, $(x\to+\infty)$. Applying Theorem 2 we see that if

$$\overline{\lim_{x\to+\infty}}\, H^*[L(x)]/x > 1, \text{ where } H^*(x) = (2/\pi)\int^{r}\log^+[T(t/k\ )]t^{-2}dt,$$

then $f(x) \equiv 0$. Making the change of variable $t = k_2u$ in equation (6) we obtain

$$H^*(r) = (2/\pi k_2)\int^{k_2r}\log^+T(u)u^{-2}du \sim k_2^{-1}H(r) \qquad (r\to+\infty),$$

by Lemma 2d. Since k_2 may be taken as near k as we please the first part of our theorem follows.

By Theorem 3 there exists a function $\phi(t)$ such that

$$|\phi(t)| \leqq e^{-\log^+ T(|t|)}(1+t^2)^{-1} \qquad (-\infty < t < \infty),$$

and such that if

$$f(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ixt}\phi(t)\,dt, \text{ then } f(x) = O\{-H^{-1}[(1-\epsilon)x]\}) \qquad (x \to +\infty)$$

for every $\epsilon > 0$. Now $f^{(n)}(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{ixt}(it)^n\phi(t)\,dt$,

$$|f^{(n)}(x)| \leqq (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} t^n[T(t)]^{-1}(1+t^2)^{-1}dt$$

$$\leqq (\pi/2)^{\frac{1}{2}} \operatorname{Max} t^n/T(t) \leqq (\pi/2)^{\frac{1}{2}} M_n{}^c \leqq (\pi/2)^{\frac{1}{2}} M_n.$$

This completes the proof for $k = 1$. For general k we need only consider $f(kx)$.

Let us agree to write $l_1(x) = \log x$, $l_2(x) = \log\log x, \cdots$, and also $e_1(x) = \exp x$, $e_2(x) = \exp\exp x, \cdots$. The following lemma is well known, but since I have been unable to find a reference the proof is included.

LEMMA 4. *If*

1. $N_n{}^{(m)} = \begin{cases} n![l_1(n)l_2(n)\cdots l_m(n)]^n & n > e_m(1) \\ n! & n \leqq e_m(1) \end{cases}$

2. $T_m(x) = \operatorname*{Max}_{n \geqq 0} x^m/N_n{}^{(m)}$,

then $\log T_m(x) \sim x[l_1(x)l_2(x)\cdots l_m(x)]^{-1}$, $(x \to +\infty)$.

Let

$$(10) \quad \log T^*{}_m(x) = \operatorname*{Max}_{n \geqq 0}[n\log x + n - nl_1(n) - nl_2(n) - \cdots - nl_m(n)].$$

Using Stirling's formula it is easily seen that if $0 < \epsilon < 1$, then, for x sufficiently large,

$$\log T^*{}_m((1-\epsilon)x) \leqq \log T_m(x) \leqq \log T^*{}_m||1+\epsilon)x),$$

so that it is sufficient to prove

$$\log T^*{}_m(x) \sim x[l_1(x)l_2(x)\cdots l_m(x)]^{-1}.$$

Let $n[x]$ be the integer for which the maximum in equation (10) is attained;

then $n[x]-1<n(x)<n[x]+1$, where $n(x)$ is the solution of the equation

$$(d/dn)[n\log x+n-nl_1(n)-\cdots-nl_m(n)]=0,$$

which we may rewrite as $\log x-L(n)=0$, where

$$L(n)=l_1(n)+l_2(n)-\cdots+l_m(n)+[l_1(n)]^{-1}+[l_1(n)l_2(n)]^{-1} \\ +\cdots+[l_1(n)l_2(n)\cdots l_m(n)]^{-1}.$$

Clearly

$$\log x=l_1(n[x])+\cdots+l_m(n[x])+o(1) \qquad (x\to+\infty).$$

Inserting this in equation (10), we find that

$$\log T^*_m(x)=n[x][1+o(1)]\sim n(x) \qquad (x\to+\infty).$$

It may be verified that if $0<\epsilon<1$, and if x is sufficiently large, then $L(n)>\log x$ when $n=(1+\epsilon)x[l_1(x)l_2(x)\cdots l_m(x)]^{-1}$, and $L(n)<\log x$ when $n=(1-\epsilon)x[l_1(x)l_2(x)\cdots l_m(x)]^{-1}$. It follows that

$$(11) \qquad n(x)\sim x[l_1(x)l_2(x)\cdots l_m(x)]^{-1} \qquad (x\to+\infty),$$

and our lemma is proved.

Let us set $H_m(r)=(2/\pi)\int^r\log T_m(u)u^{-2}du$.

By Lemma 4 we have

$$(12) \qquad H_m(r)\sim(2/\pi)l_{m+1}(r) \qquad (r\to+\infty).$$

Combining equation (12) with Theorem 4 we obtain

COROLLARY 4. *Let*

$$N_n^{(m)}=\begin{cases} n![l_1(n)l_2(n)\cdots l_m(n)]^n & n>e_m(1) \\ n! & n\leqq e_m(1) \end{cases}.$$

If

$$(13) \qquad |f^{(n)}(x)|\leqq AN_n^{(m)}(2/\pi)^n \qquad (-\infty<x<\infty;\ n=0,1,\cdots),$$

and if

$$(14) \qquad f(x)=O(e_1\{-e_{m+1}[(1+\epsilon)x]\}) \qquad (x\to+\infty)$$

for $\epsilon>0$, then $f(x)\equiv0$. On the other hand there exists a function

$f(x) \not\equiv 0$, satisfying inequalities (13) and such that relation (14) holds for every $\epsilon < 0$.

HARVARD UNIVERSITY AND WASHINGTON UNIVERSITY.

BIBLIOGRAPHY.

1. L. Bieberbach, *Lehrbuch der Funktionentheorie*, Leipzig, 1931.

2. A. Gorny, "Contribution à l'étude des fonctions dérivables d'une variable réelle," *Acta Mathematica*, vol. 71 (1939), pp. 317-358.

3. G. H. Hardy, "A theorem concerning Fourier transforms," *Journal of the London Mathematical Society*, vol. 8 (1933), pp. 227-231.

4. I. I. Hirschman, Jr., "The behaviour at infinity of certain convolution transforms," submitted to the *Transactions of the American Mathematical Society*.

5. A. Kolmogoroff, "Une généralization de l'inégalité de M. J. Hadamard entre les bornes supérieures des dérivées successives d'une fonction," *Comptes Rendus*, vol. 207 (1938), p. 764.

6. A. E. Ingham, "A note on Fourier transforms," *Journal of the London Mathematical Society*, vol. 9 (1934), pp. 29-32.

7. N. Levinson, *Gap and density theorems*, New York, 1940, pp. 73-87.

8. S. Mandelbrojt, "Analytic functions and classes of infinitely differentiable functions," *The Rice Institute Pamphlet*, vol. 29 (1942), pp. 1-142.

9. G. W. Morgan, "A note on Fourier transforms," *Journal of the London Mathematical Society*, vol. 9 (1934), pp. 187-192.

10. E. C. Titchmarsh, *The theory of functions*, Oxford, 1932.

THE MARRIAGE PROBLEM.*

By Paul R. Halmos and Herbert E. Vaughan.

In a recent issue of this journal Weyl [1] proved a combinatorial lemma which was apparently considered first by P. Hall.[2] Subsequently Everett and Whaples [3] published another proof and a generalization of the same lemma. Their proof of the generalization appears to duplicate the usual proof of Tychonoff's theorem.[4] The purpose of this note is to simplify the presentation by employing the statement rather than the proof of that result. At the same time we present a somewhat simpler proof of the original Hall lemma.

Suppose that each of a (possibly infinite) set of boys is acquainted with a finite set of girls. Under what conditions is it possible for each boy to marry one of his acquaintances? It is clearly necessary that every finite set of k boys be, collectively, acquainted with at least k girls; the Everett-Whaples result is that this condition is also sufficient.

We treat first the case (considered by Hall) in which the number of boys is finite, say n, and proceed by induction. For $n = 1$ the result is trivial. If $n > 1$ and if it happens that every set of k boys, $1 \leqq k < n$, has at least $k + 1$ acquaintances, then an arbitrary one of the boys may marry any one of his acquaintances and refer the others to the induction hypothesis. If, on the other hand, some group of k boys, $1 \leqq k < n$, has exactly k acquaintances, then this set of k may be married off by induction and, we assert, the remaining $n - k$ boys satisfy the necessary condition with respect to the as yet unmarried girls. Indeed if $1 \leqq h \leqq n - k$, and if some set of h bachelors were to know fewer than h spinsters, then this set of h bachelors together with the k married men would have known fewer than $k + h$ girls. An

* Received June 6, 1949.

[1] H. Weyl, "Almost periodic invariant vector sets in a metric vector space," *American Journal of Mathematics*, vol. 71 (1949), pp. 178-205.

[2] P. Hall, "On representation of subsets," *Journal of the London Mathematical Society*, vol. 10 (1935), pp. 26-30.

[3] C. J. Everett and G. Whaples, "Representations of sequences of sets," *American Journal of Mathematics*, vol. 71 (1949), pp. 287-293. Cf. also M. Hall, "Distinct representatives of subsets," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 922-926.

[4] C. Chevalley and O. Frink, Jr., "Bicompactness of Cartesian products," *Bulletin of the American Mathematical Society*, vol. 47 (1941), pp. 612-614.

application of the induction hypothesis to the $n-k$ bachelors concludes the proof in the finite case.

If the set B of boys is infinite, consider for each b in B the set $G(b)$ of his acquaintances, topologized by the discrete topology, so that $G(b)$ is a compact Hausdorff space. Write G for the topological Cartesian product of all $G(b)$; by Tychonoff's theorem G is compact. If $\{b_1, \cdots, b_n\}$ is any finite set of boys, consider the set H of all those elements $g = g(b)$ of G for which $g(b_i) \neq g(b_j)$ whenever $b_i \neq b_j$, $i, j = 1, \cdots, n$. The set H is a closed subset of G and, by the result for the finite case, H is not empty. Since a finite union of finite sets is finite, it follows that the class of all sets such as H has the finite intersection property and, consequently, has a non empty intersection. Since an element $g = g(b)$ in this intersection is such that $g(b') \neq g(b'')$ whenever $b' \neq b''$, the proof is complete.

It is perhaps worth remarking that this theorem furnishes the solution of the celebrated problem of the monks.[5] Without entering into the history of this well-known problem, we state it and its solution in the language of the preceding discussion. A necessary and sufficient condition that each boy b may establish a harem consisting of $n(b)$ of his acqaintances, $n(b) = 1, 2, 3, \cdots$, is that, for every finite subset B_0 of B, the total number of acquaintances of the members of B_0 be at least equal to $\Sigma n(b)$, where the summation runs over every b in B_0. The proof of this seemingly more general assertion may be based on the device of replacing each b in B by $n(b)$ replicas seeking conventional marriages, with the understanding that each replica of b is acquainted with exactly the same girls as b. Since the stated restriction on the function n implies that the replicas satisfy the Hall condition, an application of the Everett-Whaples theorem yields the desired result.

University of Chicago
and
University of Illinois.

[5] H. Balzac, *Les Cent Contes Drôlatiques*, IV, 9: *Des moines et novices*, Paris (1849).

NOTE ON A RESULT OF L. FUCHS ON ORDERED GROUPS.*[1]

By C. J. Everett.

Let G be an abelian group. If it admits a linear order, every non-zero element satisfies the relation $a > 0$ or $a < 0$, whence $na > 0$ or $na < 0$ for all $n = 1, 2, \cdots$. Hence, G has the property (*) : every non-zero element of G is of infinite order.

If P_0 is an arbitrary partial order on an abelian group G of type (*), it possesses a linear extension L. To see this, it is convenient to recall that a partial order on G is completely defined by its set N of elements $p \geqq 0$. The latter has the characterizing properties: A) N is closed under addition, B) contains zero, C) contains no element along with its inverse except zero. Note that if positive multiples nx and $m(-x)$ are in N, then[2] mnx and $nm(-x) = -(mnx)$ are in N and $x = 0$.

If N is the non-negative set of a partial order P and neither x nor $-x$ is in N, three mutually exclusive cases may obtain: 1) some positive multiple nx is in N, 2) some positive multiple $m(-x)$ is in N, 3) no positive multiple nx nor $m(-x)$ is in N. In case 1, define N^* as the set of all elements of form $p + nx$, p in N, $n = 0, 1, 2, \cdots$; in case 2, similarly but with x replaced by $-x$; in case 3, in either way. It is trivial that N^* satisfies A, B, C, and contains N properly.

The class of all sets N satisfying A, B, C, and containing the set N_0 of the original P_0, is a partially ordered set under set inclusion; and every linearly ordered subclass has an upper bound, namely its set union. Hence there are maximal sets containing N_0. But a maximal set must contain either x or $-x$ for every x of G, and the corresponding order is a linear extension of P_0.

Los Alamos Scientific Laboratory.

* Received March 16, 1949.

[1] L. Fuchs, "On the extension of the partial order of groups," *American Journal of Mathematics*, vol. 72 (1950), pp. 191-194.

[2] Remark by L. Fuchs simplifying original argument.